



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 97723

TO: Lisa V Cook
Location: cm-1/7b17/7e12
Art Unit : 1641
Thursday, July 10, 2003

Case Serial Number: 851058

From : Susan Hanley
Location: Biotech-Chem Library
CM1 6B05
Phone: 305-4053

susan.hanley@uspto.gov

Search Notes

This Page Blank (uspto)

Requester's Full Name: Lisa V. Cook Examiner #: 77134 Date: 6/30/03
Art Unit: 1641 Phone Number: 305-0808 Serial Number: 09/851058
Mail Box and Bldg/Room Location: CM1-7B-17 Results Format Preferred (circle): PAPER DISK 5-MA

Office CM1 7B-17
If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Process for analyzing protein samples
Inventors (please provide full names): Kenneth C. Parker, Timothy K. Nadler, George J. Yella, Rudolf Abersold, Marcus Smolka

Earliest Priority Filing Date: 5/8/01

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Sequence Search including interference search.

Seq ID NO: 1
Seq ID NO: 2
Seq ID NO: 3

> all three - EST's
labeled peptides.
also see attached claims and bib sheet.

seq 1 10 } N/A
2 5 }
3 5 }
Thanks @
Lisa C.

RECEIVED
JUN 30 2003
STIC

STAFF USE ONLY		Type of Search	Vendors and cost where applicable
Searcher: <u>Hanley</u>	NA Sequence (#)	STN	
Searcher Phone #:	AA Sequence (#) <u>3</u>	Dialog	
Searcher Location:	Structure (#)	Questel/Orbit	
Date Searcher Picked Up: <u>7/8</u>	Bibliographic	Dr. Link	
Date Completed: <u>7/10</u>	Litigation	Lexis/Nexis	
Searcher Prep & Review Time: <u>3</u>	Fulltext	Sequence Systems <u>absso2</u>	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:55:35 ; Search time 44 Seconds
(without alignments)
30.284 Million cell updates/sec

Title: US-09-851-058-1
Perfect score: 54
Sequence: 1 OLPCPAELLR 10

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	476	AAU29351	Novel mar regulate
2	45	83.3	79	AAU59483	Protonibacterium
3	41	75.9	50	AAM88485	Human immune/haema
4	41	75.9	78	AAAB54300	Human pancreatic c
5	41	75.9	180	AAAB43728	Human cancer assoc
6	41	75.9	445	AAAB25768	Human MLN 64, Hom
7	39	72.2	20	AAAB13666	C. trachomatis OMC
8	39	72.2	20	AAAB13667	C. trachomatis OMC
9	39	72.2	20	AAAB13667	Chlamydia trachoma
10	39	72.2	20	AAAB13667	Chlamydia trachoma

11	39	72.2	20	23	ABB94205
12	39	72.2	20	23	ABB94206
13	39	72.2	51	20	AAV12769
14	39	72.2	132	22	AAE03341
15	39	72.2	132	22	ABG64431
16	39	72.2	143	22	AAE03308
17	39	72.2	143	22	ABG64434
18	39	72.2	234	19	AAAB3973
19	39	72.2	252	15	ABB04126
20	39	72.2	553	23	AAV37768
21	39	72.2	553	23	ABB94288
22	39	72.2	554	23	AAE20293
23	38	70.4	41	22	ABB39537
24	38	70.4	41	22	AAAB60233
25	38	70.4	41	22	AAAB72858
26	38	70.4	41	22	AAAB33089
27	38	70.4	41	23	ABG42668
28	37	68.5	68	22	AAU40481
29	37	68.5	87	22	AAAB5577
30	37	68.5	87	22	AAAB78218
31	37	68.5	115	21	AAAB32866
32	37	68.5	284	22	AAU36185
33	37	68.5	405	22	ABAB7386
34	37	68.5	420	22	AAAB79310
35	37	68.5	420	23	ABAB0836
36	37	68.5	1004	22	AAU28066
37	37	68.5	1025	22	AAU28254
38	37	68.5	1025	22	AAU31447
39	37	68.5	1376	22	ABG28780
40	36	66.7	76	22	AAO04418
41	36	66.7	106	19	AAAB74752
42	36	66.7	178	22	AAAB4911
43	36	66.7	262	22	AAU41650
44	36	66.7	890	22	ABAB7223
45	35	64.8	37	22	ABG11041

ALIGNMENTS

RESULT 1	AAU29351	standard; Protein: 476 AA.
AC	AAU29351;	
DT	18-DEC-2001 (first entry)	
DE	Novel mar regulated protein (NMR) #23.	
XX	mar regulated polypeptide; NMR; microbial infection; antibacterial.	
XX	Escherichia coli.	
OS	WO200170776-A2.	
PN	27-SEP-2001.	
XX		
XX	08-MAR-2001; 2001WO-US07478.	
PF	10-MAR-2000; 2000US-188362P.	
XX		
XX	(TUFT) TUFTS COLLEGE.	
PA	Levy SB, Barbosa TM, Aleksun MN;	
XX	WPI: 2001-602769/68.	
PI	N-PSDB; AAS46251.	
XX		
DR	Identifying compounds that modulate a newly identified mar regulated	
XX	polypeptide activity, useful as antimicrobial compounds, involves	
PT	contacting the polypeptide with a test compound.	
XX		

Chlamydia peptide
Chlamydia peptide
Human 5' EST seque
Human gene 17 enco
Human albumin fusi
Human gene 17 enco
Human albumin fusi
Pseudomonas fluore
Pseudomonas fluore
Chlamydia trachoma
Chlamydia trachoma
Chlamydia trachoma
Peptide #7043 enco
Human bone marrow
Peptide #7126 enco
Human peptide enco
Protonibacterium
Human brain expres
Human bone marrow
Zea mays protein f
Pseudomonas aerugi
Drosophila melanog
Corynebacterium g1
Corynebacterium g1
Novel human secret
Novel human secret
Novel human secret
Novel human secret
Human polypeptide
Human secreted pro
Human immune/haema
Protonibacterium
Drosophila melanog
Novel human diagno

PS Disclosure; Page 302-303; 526pp; English.

XX
CC The invention relates to a method of identifying compounds that modulate
CC a newly identified mar regulated (NIMR) polypeptide activity. The method
CC comprises contacting an NIMR polypeptide with a test compound under
CC interaction conditions, determining the ability of the compound to
CC modulate the activity or expression of the polypeptide, and selecting the
CC modulators. NIMR nucleic acids and polypeptides are used in the treatment
CC of microbial infections, and in screening for modulators of NIMR
CC expression and activity. These modulators can be used to reduce the
CC infectivity of a microbe on a surface, and the virulence of a microbe in
CC a subject suffering from an infection. AAU29329-AAU29379 represent
CC Escherichia coli NIMR amino acid sequences of the invention.

XX Sequence 476 AA;

SO Query Match 100.0%; Score 54; DB 22; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLRCPAELLR 10
Db 415 QLRCPAELLR 424

RESULT 2

AAU59483
ID AAU59483 standard; Protein: 79 AA.

XX
AC AAU59483;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #20379.

XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX
OS Propionibacterium acnes.

XX
PN WO200181581-A2.

XX
PD 01-NOV-2001.

XX
PE 20-APR-2001; 2001WO-US012865.

XX
PR 21-APR-2000; 2000US-199047P.

XX
PR 02-JUN-2000; 2000US-208841P.

XX
PR 07-JUL-2000; 2000US-216747P.

XX
PA (CORI-) CORIXA CORP.

XX
PI Skelly YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX
PI L'Abbe-Duval J, Zhang Y, Jen S, Carter D;

XX
DR N-PSDB; AAS59602.

XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX
PS Example 1; SEQ ID No 20678; 1069pp; English.

XX
CC Sequences AAU9105-AAU6017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 79 AA;

SO Query Match 83.3%; Score 45; DB 22; Length 79;
Best Local Similarity 88.9%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPCPAELLR 10
Db 54 LPCPAELLR 62

RESULT 3

AAH88485
ID AAH88485 standard; Protein: 50 AA.

XX
AC AAH88485;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen SEQ ID NO:16078.

XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.

XX
OS Homo sapiens.

XX
PN WO200157182-A2.

XX
PD 09-AUG-2001.

XX
PE 17-JAN-2001; 2001WO-US01354.

XX
PR 31-JAN-2000; 2000US-0179065.

XX
PR 04-FEB-2000; 2000US-0180628.

XX
PR 24-FEB-2000; 2000US-0184664.

XX
PR 16-MAR-2000; 2000US-0186350.

XX
PR 17-MAR-2000; 2000US-0189874.

XX
PR 18-APR-2000; 2000US-0190076.

XX
PR 19-MAY-2000; 2000US-0205515.

XX
PR 07-JUN-2000; 2000US-0209467.

XX
PR 28-JUN-2000; 2000US-0214886.

XX
PR 30-JUN-2000; 2000US-0215135.

XX
PR 07-JUL-2000; 2000US-0216647.

XX
PR 11-JUL-2000; 2000US-0216880.

XX
PR 11-JUL-2000; 2000US-0217487.

XX
PR 14-JUL-2000; 2000US-0218290.

XX
PR 26-JUL-2000; 2000US-0220963.

XX
PR 14-AUG-2000; 2000US-0224518.

XX
PR 14-AUG-2000; 2000US-0224519.

XX
PR 14-AUG-2000; 2000US-0225214.

XX
PR 14-AUG-2000; 2000US-0225266.

XX
PR 14-AUG-2000; 2000US-0225267.

XX
PR 14-AUG-2000; 2000US-0225268.

XX
PR 14-AUG-2000; 2000US-0225270.

XX
PR 14-AUG-2000; 2000US-0225447.

OY 1 QLPCPAE 7
 DB 14 QLPCPAE 20
 Best Local Similarity 87.5%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAELL 9
 DB 54 LPCPAELV 61

RESULT 4
 AAB54300
 ID AAB54300 standard; Protein: 78 AA.

AC AAB54300;
 XX
 DT 09-MAR-2001 (first entry)

DE Human pancreatic cancer antigen protein sequence SEQ ID NO:752.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW neotropic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.

XX Homo sapiens.
 OS
 PN WO200055320-A1.

XX 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05989.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI: 2000-579444/54.

DR N-PSDB: AAC99065.

PT New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 XX
 PS Claim 11; Page 1195-1196; 1379PP; English.

CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, neotropic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiant and antiinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.

XX Sequence 78 AA;

Query Match 75.9%; Score 41; DB 21; Length 78;

RESULT 5
 AAB43728
 ID AAB43728 standard; Protein: 180 AA.

AC AAB43728;
 XX
 DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1173.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
 KW antidiabetic; antiaesthetic; antirheumatic; antitarrtic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW allergic reaction; haematopoietic cell disorder; autoimmune disorder;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

XX Homo sapiens.
 OS

PN WO200055350-A1.

XX 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05882.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI: 2000-587533/55.

DR N-PSDB: AAC77937.

PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 11; Page 1794-1795; 2352PP; English.

CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerrary; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antitarrtic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

SQ Sequence 180 AA;
Query Match 75.98; Score 41; DB 21; Length 180;
Best Local Similarity 87.58; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 LPCPAELV 9
| | | | | | | | | |
Db 24 LPCPAELV 31

RESULT 6
AAW25768
ID AAW25768 standard; Protein; 445 AA.
XX
AC AAW25768;
XX
DT 05-DEC-1997 (first entry)
XX
DE Human MLN 64.
XX
XX Human; CAR11; Lasp-1; MLN 64; hD53; MD52; breast carcinoma cell;
XX Cys-rich region; RING finger domain; CAR motif; TRAF domain; LIM domain;
XX TNF-related cytokine signal transduction; breast cancer; SH3;
XX Src homology region 3.
XX
OS Homo sapiens.
XX
FH
FT Domain Location/Qualifiers
FT 1..72 /label= Potential_Transmembrane_domain
FT 94..168 /label= Potential_Transmembrane_domain
FT Domain
XX
XX MO9706256-A2.
XX
XX 20-FEB-1997.
XX
XX 31-JUL-1996; 96MO-US12500.
XX
XX 09-AUG-1995; 95US-0002183.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX (CNRS) CENT NAT RECH SCI.
XX (INRM) INST NAT SANTE & RECH MEDICALE.
XX (UYPA-) UNIV PASTEUR LOUIS.
XX
XX Bassett P, Byrne J, Rio M, Tomasetto C;
XX
XX WPI; 1997-154263/14.
XX N-PSDB; AAT86171.
XX
XX Genes amplified and over-expressed in breast carcinoma - located on
XX chromosome 17 q11-q21.3 or 6 q22-q23, useful in breast cancer
XX prognosis and as leukemia markers
XX
XX Claim 11: Fig 16; 197pp; English.

This sequence is human MLN 64 encoded by DNA deposited as ATCC 97609.
The MLN 64 gene has been mapped to the q12-q21 region of the long arm
of chromosome 17 with the maximum in the q21.1. This region already
contains two genes known to be involved in breast cancer. The MLN 64
gene contains 15 exons. There are 7 distinct MLN 64 cDNAs resulting
from nucleotide substitutions, deletions and/or insertions. These
modifications mainly occur at Intron/exon boundaries, suggesting that
the MLN 64 variants result from defective splicing processes. Two
C-truncated MLN 64 proteins. Also, 3 of these lead to chimeric proteins
containing additive nonsense protein sequences of 16, 20 and 126
residues. The full length MLN 64 protein has two potential
transmembrane domains and several potential leucine zipper and
leucine-rich repeat structures. The MLN 64 gene and protein are
useful as prognostic markers for breast cancer.

XX
SQ Sequence 445 AA;
Query Match 75.98; Score 41; DB 18; Length 445;
Best Local Similarity 87.58; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 LPCPAELV 9
| | | | | | | | | |
Db 289 LPCPAELV 296

RESULT 7
AAB13666
ID AAB13666 standard; Peptide; 20 AA.
XX
AC AAB13666;
XX
DT 02-FEB-2001 (first entry)
XX
DE C. trachomatis OMCB peptide 133-152.
XX
XX Chlamydial infection; sexually transmitted disease;
XX pelvic inflammatory disease; PID; tubal obstruction; infertility;
XX trachoma; blindness; acute respiratory tract infection;
XX atherosclerosis; coronary heart disease; antibacterial.
XX
XX Chlamydia trachomatis.
XX
XX
XX WO200034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99MO-US29012.
XX
XX 08-DEC-1998; 98US-0208277.
XX 08-APR-1999; 99US-0288594.
XX 01-OCT-1999; 99US-0410568.
XX 22-OCT-1999; 99US-0426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises
XX amino acid sequence encoded by polynucleotide sequence
XX
XX Disclosure: Page 231; 256pp; English.

The present invention relates to new nucleic acid sequences and the
proteins encoded by the nucleic acid sequences. The encoded proteins
comprise an immunogenic portion of a Chlamydia antigen. The encoded
proteins are useful for the serodiagnosis and treatment of Chlamydia
infection. Chlamydiae are intracellular bacterial pathogens that are
responsible for a wide variety of human infections. C. trachomatis
infection is one of the most common sexually transmitted diseases and can
lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
and infertility. Trachoma due to ocular infection with C. trachomatis is
the leading cause of preventable blindness worldwide. C. pneumoniae is a
major cause of acute respiratory tract infections in humans and is also
thought to play a role in the pathogenesis of atherosclerosis and
coronary heart disease. The present sequence is a protein isolated in the
present invention.

SQ Sequence 20 AA;
Query Match 72.28; Score 39; DB 21; Length 20;
Best Local Similarity 70.08; Pred. No. 9.6;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCPAELLR 10
 ||||| : :
 Db 7 QLPCBAEFVR 16

RESULT 8

ID AAB13667 standard; Peptide: 20 AA.
 AC AAB13667;

DT 02-FEB-2001 (first entry)
 DE C. trachomatis OMCB peptide 137-156.

XX Chlamydia infection; sexually transmitted disease;
 KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
 KW trachoma; blindness; acute respiratory tract infection;
 KW atherosclerosis; coronary heart disease; antibacterial.

OS Chlamydia trachomatis.

PN WO200034483-A2.

PD 15-JUN-2000.

PF 08-DEC-1999; 99WO-US29012.

PR 08-DEC-1998; 98US-0208277.

PR 08-APR-1999; 99US-0288594.

PR 01-OCT-1999; 99US-0410568.

PR 22-OCT-1999; 99US-0426571.

PA (CORI-) CORIXA CORP.

PI Probst P, Bhatia A, Skelky YAW, Fling SP, Jen S, Stromberg EJ;
 DR WPI: 2000-431303/37.

XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection

PT comprises immunogenic portion of Chlamydia antigen, which comprises
 amino acid sequence encoded by polynucleotide sequence -

XX Disclousure: Page 231; 256pp; English.

XX The present invention relates to new nucleic acid sequences and the
 CC proteins encoded by the nucleic acid sequences. The encoded proteins
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia
 CC infection. Chlamydiae are intracellular bacterial pathogens that are
 CC responsible for a wide variety of human infections. C. trachomatis
 CC infection is one of the most common sexually transmitted diseases and can
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is
 CC the leading cause of preventable blindness worldwide. C. pneumoniae is a
 CC major cause of acute respiratory tract infections in humans and is also
 CC thought to play a role in the pathogenesis of atherosclerosis and
 CC coronary heart disease. The present sequence is a protein isolated in the
 CC present invention.

XX Sequence 20 AA;

Query Match 72.2%; Score 39; DB 21; Length 20;
 Best Local Similarity 70.0%; Pred. No. 9.6;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCPAELLR 10
 ||||| : :
 Db 3 QLPCBAEFVR 12

RESULT 9
 AAG83234

ID AAG83234 standard; Peptide: 20 AA.
 XX AAG83234;

AC AAG83234;

DT 05-SEP-2001 (first entry)
 DE Chlamydia trachomatis OMCB peptide 133-152.

XX Chlamydia; vaccine; infection; fusion protein; antigen;
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
 KW acute respiratory tract infection; Cap1; CT529; OMCB;
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.

OS Chlamydia trachomatis.

PN WO200140474-A2.

PD 07-JUN-2001.

PF 04-DEC-2000; 2000WO-US32919.

PR 03-DEC-1999; 99US-0454684.

PR 19-APR-2000; 2000US-0556877.

PR 20-JUN-2000; 2000US-0598419.

PA (CORI-) CORIXA CORP.

PI Probst P, Bhatia A, Skelky YAW, Fling SP, Scholler J;
 DR WPI: 2001-374831/39.

XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic

PT inflammatory disease, trachoma, acute respiratory tract infections,
 PI atherosclerosis and heart disease -

PS Claim 66; Page 237; 295pp; English.

XX The present peptide is provided in a specification relating to
 CC compounds and methods for the treatment and diagnosis of chlamydial
 CC infection. The compounds provided include polypeptides and fusion
 CC proteins comprising immunogenic portions of Chlamydia antigens
 CC and DNA sequences encoding such polypeptides. They are useful for
 CC vaccinating against chlamydial infection, which causes pelvic
 CC inflammatory disease, trachoma, acute respiratory tract infections,
 CC atherosclerosis and heart disease.

XX Sequence 20 AA;

Query Match 72.2%; Score 39; DB 22; Length 20;
 Best Local Similarity 70.0%; Pred. No. 9.6;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCPAELLR 10
 ||||| : :
 Db 7 QLPCBAEFVR 16

RESULT 10

ID AAG83235 standard; Peptide: 20 AA.

AC AAG83235;

DT 05-SEP-2001 (first entry)

DE Chlamydia trachomatis OMCB peptide 137-156.

XX Chlamydia; vaccine; infection; fusion protein; antigen;
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
 KW acute respiratory tract infection; Cap1; CT529; OMCB;
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
 OS Chlamydia trachomatis.

XX MO200140474-A2.
PN 07-JUN-2001.
XX
XX
XX 04-DEC-2000; 2000MO-US32919.
XX
XX 03-DEC-1999; 99US-0454684.
PR 19-APR-2000; 2000US-0556877.
PR 20-JUN-2000; 2000US-0598419.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skelky YAW, Fling SP, Scholler J;
XX WPI; 2001-374831/39.
XX
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT inflammatory disease, trachoma, acute respiratory tract infections,
PT atherosclerosis and heart disease -
XX
XX
XX Claim 66; Page 237; 295pp; English.
XX
XX The present peptide is provided in a specification relating to
CC compounds and methods for the treatment and diagnosis of chlamydial
CC infection. The compounds provided include polypeptides and fusion
CC proteins comprising immunogenic portions of Chlamydia antigens
CC and DNA sequences encoding such polypeptides. They are useful for
CC vaccinating against chlamydial infection, which causes pelvic
CC inflammatory disease, trachoma, acute respiratory tract infections,
CC atherosclerosis and heart disease.
XX
SQ Sequence 20 AA:

Query Match 72.2%; Score 39; DB 22; Length 20;
Best Local Similarity 70.0%; Pred. No. 9.6;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Caps 0;

OY 1 QLPCPALLR 10
 ||||| 11 : 1
DB 3 QLPCPAAEFVR 12

RESULT 11
ABB94205
ID ABB94205 standard; Peptide: 20 AA.
XX
AC ABB94205;
XX
DT 05-JUN-2002 (first entry)
XX
DE Chlamydia peptide sequence SEQ ID NO:242.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
KW antigen; antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response.
XX
OS Chlamydia sp.
OS Synthetic.
XX
PN WO200208267-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US23121.
XX
PR 20-JUL-2000; 2000US-0620412.
PR 23-APR-2001; 2001US-0841132.
XX
PA (CORI-) CORIXA CORP.
XX
PI Fling SP, Skelky YAW, Probst P, Bhatia A;
XX

DR WPI; 2002-179901/23.
XX
XX
PT Novel compositions comprising Chlamydia Cap1 protein and its use in the
PT treatment of Chlamydia infection -
XX
XX
XX Disclosure; Page 266; 537pp; English.
XX
XX The present invention describes compositions comprising a Chlamydia Cap1
CC protein and methods for the diagnosis and therapy of Chlamydia infection.
CC Chlamydia DNA and protein sequences from the present invention can have
CC antibacterial and immunostimulant activities, and can be used in
CC vaccines. Compounds from the present invention can be used for eliciting
CC an immune response, specifically stimulating a Chlamydia-specific T-cell
CC response or inhibiting the development of a Chlamydia infection in an
CC animal. Methods from the present invention can be used: for detecting the
CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC specific for a Chlamydia protein; and for treatment of a Chlamydia
CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 20 AA:

Query Match 72.2%; Score 39; DB 23; Length 20;
Best Local Similarity 70.0%; Pred. No. 9.6;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Caps 0;

OY 1 QLPCPALLR 10
 ||||| 11 : 1
DB 7 QLPCPAAEFVR 16

RESULT 12
ABB94206
ID ABB94206 standard; Peptide: 20 AA.
XX
AC ABB94206;
XX
DT 05-JUN-2002 (first entry)
XX
DE Chlamydia peptide sequence SEQ ID NO:243.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
KW antigen; antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response.
XX
OS Chlamydia sp.
OS Synthetic.
XX
PN WO200208267-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US23121.
XX
PR 20-JUL-2000; 2000US-0620412.
PR 23-APR-2001; 2001US-0841132.
XX
PA (CORI-) CORIXA CORP.
XX
PI Fling SP, Skelky YAW, Probst P, Bhatia A;
XX
DR WPI; 2002-179901/23.
XX
PT Novel compositions comprising Chlamydia Cap1 protein and its use in the
PT treatment of Chlamydia infection -
XX
XX Disclosure; Page 266; 537pp; English.
XX
XX The present invention describes compositions comprising a Chlamydia Cap1
CC protein and methods for the diagnosis and therapy of Chlamydia infection.
CC Chlamydia DNA and protein sequences from the present invention can have
CC antibacterial and immunostimulant activities, and can be used in
CC vaccines. Compounds from the present invention can be used for eliciting

CC an immune response, specifically stimulating a Chlamydia-specific T-cell
CC response or inhibiting the development of a Chlamydia infection in an
CC animal. Methods from the present invention can be used for detecting the
CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC specific for a Chlamydia protein; and for treatment of a Chlamydia
CC infection. AB192394 to AB192709 and AB894096 to AB894374 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 20 AA;
Query Match 72.2%; Score 39; DB 23; Length 20;
Best Local Similarity 70.0%; Pred. No. 9.6;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 QLPCEAELR 10
Db 3 QLPCEAFVR 12
RESULT 13
ID AAY12769 standard; Protein; 51 AA.
XX
XX AAY12769;
XX
XX 21-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO:359.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
XX
XX MO9906549-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98MO-IB01331.
XX
XX 01-AUG-1997; 97US-0905279.
XX
XX (GEST) GENSET.
XX
XX Duclet A, Dumas Mline Edwards J, Lacroix B;
XX
XX WPI; 1999-153779/13.
XX
XX N-PSDB; AAX51547.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries derived from testis, ovary, uterus and spleen tissue
XX
XX
XX Claim 34; Page 435-436; 522pp; English.
XX
XX AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12681 to
XX AAY12913, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for

CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX
SQ Sequence 51 AA;
Query Match 72.2%; Score 39; DB 20; Length 51;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 QLPCEAELR 9
Db 13 QLPCEPISL 21
RESULT 14
ID AAE03341 standard; Protein; 132 AA.
XX
XX AAE03341;
XX
XX 10-AUG-2001 (first entry)
XX
XX Human gene 17 encoded secreted protein H70128, SEQ ID NO:115.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angioecnic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulnerability;
XX cell culture; chemotaxis; food additive; gene therapy;
XX binding partner identification.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH 1..31
XX FT /label= Signal_peptide
XX FT 32..132
XX FT /label= Human_mature_secreted_protein
XX FT 104
XX FT /label= Unknown
XX FT /note= "Encoded by TGN"
XX FT 120
XX FT /label= Unknown
XX FT /note= "Encoded by KGG"
XX
XX WO200134800-A1.
XX
XX 17-MAY-2001.
XX
XX 08-NOV-2000; 2000MO-US10674.
XX
XX 12-NOV-1999; 99US-0164750.
XX
XX 30-JUN-2000; 2000US-0215128.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Ebner R, Fiscella M, Wei P;
XX
XX WPI; 2001-329085/34.
XX
XX N-PSDB; AAD07754.
XX
XX New nucleic acid molecules encoding human secreted proteins, used in
XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
XX Parkinson's diseases and cancers -
XX
XX Claim 11; Page 494; 530pp; English.
XX
XX AAD07705-AAD07759 represent cDNAs corresponding to 19 human secreted

PA (HUMA-) HUMAN GENOME SCI INC.

Search completed: July 9, 2003, 12:00:05
Job time : 46 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:58:36 : Search time 22.5 seconds
(Without alignments)
13.077 Million cell updates/sec

Title: US-09-851-058-1
Perfect score: 54
Sequence: 1 OLCPAELRLR 10

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_patents_AA:*
2: /cgn2_6/p/odata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/p/odata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/p/odata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/p/odata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/p/odata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	75.9	445	2	US-08-691-814B-6
2	39	72.2	20	4	US-09-556-877-242
3	39	72.2	20	4	US-09-556-877-243
4	39	72.2	20	4	US-09-620-412C-242
5	39	72.2	20	4	US-09-620-412C-243
6	39	72.2	252	1	US-08-411-777-8
7	39	72.2	252	1	US-09-057-088-8
8	36	66.7	106	4	US-09-149-476-341
9	35	64.8	234	4	US-09-230-665-6
10	35	64.8	284	3	US-08-411-777-10
11	35	64.8	284	3	US-09-057-088-10
12	35	64.8	285	4	US-09-230-665-6
13	35	64.8	286	4	US-09-234-733-3
14	35	64.8	304	4	US-09-189-060B-72
15	35	64.8	305	1	US-08-090-013-2
16	35	64.8	305	1	US-08-081-328-2
17	35	64.8	305	1	US-08-237-249-2
18	35	64.8	305	2	US-08-921-426-8
19	35	64.8	305	2	US-08-833-642A-2
20	35	64.8	305	2	US-08-140-008A-4
21	35	64.8	305	2	US-08-836-340-1
22	35	64.8	305	2	US-08-389-423-2
23	35	64.8	305	3	US-08-816-915-8
24	35	64.8	305	4	US-09-230-222-1
25	35	64.8	305	4	US-09-189-060B-56
26	35	64.8	305	4	US-09-230-665-2
27	35	64.8	305	4	US-09-189-028-2

28	35	64.8	305	5	PCT-US95-07743-8	Sequence 8, Appl1
29	35	64.8	306	4	US-09-189-060B-66	Sequence 66, Appl
30	35	64.8	306	4	US-09-189-060B-68	Sequence 68, Appl
31	35	64.8	307	4	US-09-189-060B-74	Sequence 74, Appl
32	35	64.8	308	4	US-09-189-060B-70	Sequence 70, Appl
33	35	64.8	2291	2	US-08-286-819A-29	Sequence 29, Appl
34	35	64.8	2291	3	US-08-980-357-29	Sequence 29, Appl
35	35	64.8	2618	4	US-09-413-814-28	Sequence 28, Appl
36	34	63.0	11	4	US-09-422-712B-4	Sequence 4, Appl1
37	34	63.0	11	4	US-09-607-756-4	Sequence 4, Appl1
38	34	63.0	20	4	US-09-556-877-241	Sequence 241, App
39	34	63.0	20	4	US-09-620-412C-241	Sequence 241, App
40	33	61.1	30	2	US-08-620-151-133	Sequence 133, App
41	33	61.1	299	2	US-08-872-437-2	Sequence 2, Appl1
42	33	61.1	299	3	US-08-651-136C-12	Sequence 12, Appl
43	33	61.1	299	4	US-09-229-911A-12	Sequence 12, Appl
44	33	61.1	448	4	US-09-342-681C-17	Sequence 17, Appl
45	33	61.1	448	4	US-09-342-681C-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-691-814B-6
Sequence 6, Application US/08691814B

Patent No. 5981218

GENERAL INFORMATION:

APPLICANT: RIO, Marie-Christine

APPLICANT: Tomasetto, Catherine

APPLICANT: Bassel, Paul

APPLICANT: Byrne, Jennifer

TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful

TITLE OF INVENTION: as Leukemia Markers and in Breast Cancer Prognosis

NUMBER OF SEQUENCES: 124

CORRESPONDENCE ADDRESS:

ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Ave, NW, Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/691,814B

FILING DATE: 31-JUL-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/002,183

FILING DATE: 09-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: Steffe, Eric K.

REGISTRATION NUMBER: 36,688

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2543

TELEFAX: 202-371-2543

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 445 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-691-814B-6

Query Match 75.9%, Score 41; DB 2; Length 445;
Best Local Similarity 87.5%, Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

OY 2 LCPPEALL 9
 DB 289 LCPPEALL 296

RESULT 2

US-09-556-877-242
 ; Sequence 242, Application US/09556877
 ; Patent No. 6432916
 ; GENERAL INFORMATION:
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Bhalla, Ajay
 ; APPLICANT: Skeiky, Yasir
 ; APPLICANT: Fling, Steve
 ; APPLICANT: Maisonneuve, Jeff
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C5
 ; CURRENT APPLICATION NUMBER: US/09/556,877
 ; NUMBER OF SEQ ID NOS: 305
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 242
 ; LENGTH: 20
 ; TYPE: PRF
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Made in a lab
 US-09-556-877-242

Query Match

Best Local Similarity 72.2%; Score 39; DB 4; Length 20;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 10
 DB 7 QLPCEAEFVR 16

RESULT 3

US-09-556-877-243
 ; Sequence 243, Application US/09556877
 ; Patent No. 6432916
 ; GENERAL INFORMATION:
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Bhalla, Ajay
 ; APPLICANT: Skeiky, Yasir
 ; APPLICANT: Fling, Steve
 ; APPLICANT: Maisonneuve, Jeff
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C5
 ; CURRENT APPLICATION NUMBER: US/09/556,877
 ; NUMBER OF SEQ ID NOS: 305
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 243
 ; LENGTH: 20
 ; TYPE: PRF
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Made in a lab
 US-09-556-877-243

Query Match

Best Local Similarity 72.2%; Score 39; DB 4; Length 20;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 10
 DB 3 QLPCEAEFVR 12

RESULT 4

US-09-620-412C-242
 ; Sequence 242, Application US/09620412C
 ; Patent No. 6448234
 ; GENERAL INFORMATION:
 ; APPLICANT: Steven P. Fling
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C7
 ; CURRENT APPLICATION NUMBER: US/09/620,412C
 ; NUMBER OF SEQ ID NOS: 363
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 242
 ; LENGTH: 20
 ; TYPE: PRF
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Made in a lab
 US-09-620-412C-242

Query Match

Best Local Similarity 72.2%; Score 39; DB 4; Length 20;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 10
 DB 7 QLPCEAEFVR 16

RESULT 5

US-09-620-412C-243
 ; Sequence 243, Application US/09620412C
 ; Patent No. 6448234
 ; GENERAL INFORMATION:
 ; APPLICANT: Steven P. Fling
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C7
 ; CURRENT APPLICATION NUMBER: US/09/620,412C
 ; NUMBER OF SEQ ID NOS: 363
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 243
 ; LENGTH: 20
 ; TYPE: PRF
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Made in a lab
 US-09-620-412C-243

Query Match

Best Local Similarity 72.2%; Score 39; DB 4; Length 20;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 10
 DB 3 QLPCEAEFVR 12

RESULT 6

US-08-411-777-8
 ; Sequence 8, Application US/08411777
 ; Patent No. 5792641
 ; GENERAL INFORMATION:
 ; APPLICANT: Schuelin, Martin
 ; APPLICANT: Fredholm, Henrik
 ; APPLICANT: Hjorth, Carsten
 ; APPLICANT: Rasmussen, Grethe
 ; APPLICANT: Nielsen, Egon
 ; APPLICANT: Rosholm, Peter
 ; TITLE OF INVENTION: Cellulase Variants
 ; NUMBER OF SEQUENCES: 10

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 57926410 No. 5792641disk of No. 5792641th America
;; STREET: 405 Lexington Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10174-6401
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/411,777
;; FILING DATE: 05-MAY-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gregg, Valeta A
;; REGISTRATION NUMBER: 35,127
;; REFERENCE/DOCKET NUMBER: 3913, 504-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-868-9655
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 252 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 5792641e
;; US-08-411-777-8

Query Match 72.2%; Score 39; DB 1; Length 252;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 OLPCPDEL 8
Db 226 EVPCPDEL 233

RESULT 7
US-09-057-088-8
; Sequence 8, Application US/09057088
; Patent No. 6114296
; GENERAL INFORMATION:
; APPLICANT: Schuelein, Martin
; APPLICANT: Fredholm, Henrik
; APPLICANT: Hjorth, Carsten
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Nielsen, Egon
; APPLICANT: Rosholm, Peter
; TITLE OF INVENTION: Cellulase Variants
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61142960 No. 6114296disk of No. 6114296th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,088
; FILING DATE:

;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/411,777
;; FILING DATE: 05-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gregg, Valeta A
;; REGISTRATION NUMBER: 35,127
;; REFERENCE/DOCKET NUMBER: 3913, 504-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-868-9655
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 252 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 6114296e
;; US-09-057-088-8

Query Match 72.2%; Score 39; DB 3; Length 252;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 OLPCPDEL 8
Db 226 EVPCPDEL 233

RESULT 8
US-09-149-476-341
; Sequence 341, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503

1	EARLIER	FILING DATE:	1997-05-23
2	EARLIER	APPLICATION NUMBER:	60/047, 599
3	EARLIER	FILING DATE:	1997-05-23
4	EARLIER	APPLICATION NUMBER:	60/047, 581
5	EARLIER	FILING DATE:	1997-05-23
6	EARLIER	APPLICATION NUMBER:	60/047, 584
7	EARLIER	FILING DATE:	1997-05-23
8	EARLIER	APPLICATION NUMBER:	60/047, 500
9	EARLIER	FILING DATE:	1997-05-23
10	EARLIER	APPLICATION NUMBER:	60/047, 587
11	EARLIER	FILING DATE:	1997-05-23
12	EARLIER	APPLICATION NUMBER:	60/047, 592
13	EARLIER	FILING DATE:	1997-05-23
14	EARLIER	APPLICATION NUMBER:	60/047, 596
15	EARLIER	FILING DATE:	1997-05-23
16	EARLIER	APPLICATION NUMBER:	60/047, 611
17	EARLIER	FILING DATE:	1997-05-23
18	EARLIER	APPLICATION NUMBER:	60/047, 632
19	EARLIER	FILING DATE:	1997-05-23
20	EARLIER	APPLICATION NUMBER:	60/047, 601
21	EARLIER	FILING DATE:	1997-05-23
22	EARLIER	APPLICATION NUMBER:	60/043, 580
23	EARLIER	FILING DATE:	1997-04-11
24	EARLIER	APPLICATION NUMBER:	60/043, 566
25	EARLIER	FILING DATE:	1997-04-11
26	EARLIER	APPLICATION NUMBER:	60/043, 614
27	EARLIER	FILING DATE:	1997-04-11
28	EARLIER	APPLICATION NUMBER:	60/043, 569
29	EARLIER	FILING DATE:	1997-04-11
30	EARLIER	APPLICATION NUMBER:	60/043, 611
31	EARLIER	FILING DATE:	1997-04-11
32	EARLIER	APPLICATION NUMBER:	60/043, 617
33	EARLIER	FILING DATE:	1997-04-11
34	EARLIER	APPLICATION NUMBER:	60/043, 674
35	EARLIER	FILING DATE:	1997-04-11
36	EARLIER	APPLICATION NUMBER:	60/043, 669
37	EARLIER	FILING DATE:	1997-04-11
38	EARLIER	APPLICATION NUMBER:	60/043, 612
39	EARLIER	FILING DATE:	1997-04-11
40	EARLIER	APPLICATION NUMBER:	60/043, 613
41	EARLIER	FILING DATE:	1997-04-11
42	EARLIER	APPLICATION NUMBER:	60/043, 672
43	EARLIER	FILING DATE:	1997-04-11
44	EARLIER	APPLICATION NUMBER:	60/043, 672
45	EARLIER	FILING DATE:	1997-04-11
46	EARLIER	APPLICATION NUMBER:	60/043, 615
47	EARLIER	FILING DATE:	1997-04-11
48	EARLIER	APPLICATION NUMBER:	60/048, 974
49	EARLIER	FILING DATE:	1997-06-06
50	EARLIER	APPLICATION NUMBER:	60/056, 886
51	EARLIER	FILING DATE:	1997-08-22
52	EARLIER	APPLICATION NUMBER:	60/056, 877
53	EARLIER	FILING DATE:	1997-08-22
54	EARLIER	APPLICATION NUMBER:	60/056, 889
55	EARLIER	FILING DATE:	1997-08-22
56	EARLIER	APPLICATION NUMBER:	60/056, 893
57	EARLIER	FILING DATE:	1997-08-22
58	EARLIER	APPLICATION NUMBER:	60/056, 630
59	EARLIER	FILING DATE:	1997-08-22
60	EARLIER	APPLICATION NUMBER:	60/056, 878
61	EARLIER	FILING DATE:	1997-08-22
62	EARLIER	APPLICATION NUMBER:	60/056, 662
63	EARLIER	FILING DATE:	1997-08-22
64	EARLIER	APPLICATION NUMBER:	60/056, 872
65	EARLIER	FILING DATE:	1997-08-22
66	EARLIER	APPLICATION NUMBER:	60/056, 882
67	EARLIER	FILING DATE:	1997-08-22
68	EARLIER	APPLICATION NUMBER:	60/056, 637
69	EARLIER	FILING DATE:	1997-08-22

1	EARLIER APPLICATION NUMBER: 60/056,903
2	EARLIER FILING DATE: 1997-08-22
3	EARLIER APPLICATION NUMBER: 60/056,888
4	EARLIER FILING DATE: 1997-08-22
5	EARLIER APPLICATION NUMBER: 60/056,879
6	EARLIER FILING DATE: 1997-08-22
7	EARLIER APPLICATION NUMBER: 60/056,864
8	EARLIER FILING DATE: 1997-08-22
9	EARLIER APPLICATION NUMBER: 60/056,858
10	EARLIER FILING DATE: 1997-08-22
11	EARLIER APPLICATION NUMBER: 60/056,849
12	EARLIER FILING DATE: 1997-08-22
13	EARLIER APPLICATION NUMBER: 60/056,845
14	EARLIER FILING DATE: 1997-08-22
15	EARLIER APPLICATION NUMBER: 60/056,832
16	EARLIER FILING DATE: 1997-08-22
17	EARLIER APPLICATION NUMBER: 60/057,761
18	EARLIER FILING DATE: 1997-08-22
19	EARLIER APPLICATION NUMBER: 60/047,585
20	EARLIER FILING DATE: 1997-05-23
21	EARLIER APPLICATION NUMBER: 60/047,586
22	EARLIER FILING DATE: 1997-05-23
23	EARLIER APPLICATION NUMBER: 60/047,550
24	EARLIER FILING DATE: 1997-05-23
25	EARLIER APPLICATION NUMBER: 60/047,594
26	EARLIER FILING DATE: 1997-05-23
27	EARLIER APPLICATION NUMBER: 60/047,589
28	EARLIER FILING DATE: 1997-05-23
29	EARLIER APPLICATION NUMBER: 60/047,593
30	EARLIER FILING DATE: 1997-05-23
31	EARLIER APPLICATION NUMBER: 60/047,614
32	EARLIER FILING DATE: 1997-05-23
33	EARLIER APPLICATION NUMBER: 60/043,670
34	EARLIER FILING DATE: 1997-04-11
35	EARLIER APPLICATION NUMBER: 60/043,578
36	EARLIER FILING DATE: 1997-04-11
37	EARLIER APPLICATION NUMBER: 60/043,576
38	EARLIER FILING DATE: 1997-04-11
39	EARLIER APPLICATION NUMBER: 60/047,501
40	EARLIER FILING DATE: 1997-05-23
41	EARLIER APPLICATION NUMBER: 60/043,670
42	EARLIER FILING DATE: 1997-04-11
43	EARLIER APPLICATION NUMBER: 60/056,632
44	EARLIER FILING DATE: 1997-08-22
45	EARLIER APPLICATION NUMBER: 60/056,664
46	EARLIER FILING DATE: 1997-08-22
47	EARLIER APPLICATION NUMBER: 60/056,876
48	EARLIER FILING DATE: 1997-08-22
49	EARLIER APPLICATION NUMBER: 60/056,881
50	EARLIER FILING DATE: 1997-08-22
51	EARLIER APPLICATION NUMBER: 60/056,909
52	EARLIER FILING DATE: 1997-08-22
53	EARLIER APPLICATION NUMBER: 60/056,875
54	EARLIER FILING DATE: 1997-08-22
55	EARLIER APPLICATION NUMBER: 60/056,862
56	EARLIER FILING DATE: 1997-08-22
57	EARLIER APPLICATION NUMBER: 60/056,887
58	EARLIER FILING DATE: 1997-08-22
59	EARLIER APPLICATION NUMBER: 60/056,908
60	EARLIER FILING DATE: 1997-08-22

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 66.7%; Score 36; DB 4; Length 106;
Best Local Similarity 57.1%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 0; Indels 4; Gaps 1;

OY 1 QLPCP---AELLR 10
| | | | | : | | | :
Db 36 QLPCPLLSAQLK 49

RESULT 9
US-09-230-665-6
Sequence 6, Application US/09230665
Patent No. 6322595
GENERAL INFORMATION:
APPLICANT: Boyer, Stanton L
TITLE OF INVENTION: Detergent Composition Comprising Two Cellulase
FILE REFERENCE: 6191 Sequence Listing (8 Sequences)
Patent No. 6322595
CURRENT APPLICATION NUMBER: US/09/230,665
CURRENT FILING DATE: 1999-01-29
EARLIER APPLICATION NUMBER: 60/023,125
EARLIER FILING DATE: 1996-07-30
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 234
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Organism encoding
OTHER INFORMATION: an endoglucanase enzyme
US-09-230-665-6

Query Match 64.8%; Score 35; DB 4; Length 234;
Best Local Similarity 66.7%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QLPCPAELL 9
| : | | | | :
Db 207 QVOCPAELV 215

RESULT 10
US-08-411-777-10
Sequence 10, Application US/08411777
Patent No. 5792641
GENERAL INFORMATION:
APPLICANT: Schuelein, Martin
APPLICANT: Fredholm, Henrik
APPLICANT: Hjorth, Carsten
APPLICANT: Rasmussen, Grethe
APPLICANT: Nielsen, Egon
APPLICANT: Rosholm, Peter
TITLE OF INVENTION: Cellulase Variants
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5792641o No. 5792641disk of No. 5792641th America
STREET: 405 Lexington Avenue

CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/411,777
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeria A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 3913,504-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-868-9655
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 284 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5792641e
US-08-411-777-10

Query Match 64.8%; Score 35; DB 1; Length 284;
Best Local Similarity 66.7%; Pred. No. 11e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QLPCPAELL 9
| : | | | | :
Db 186 QVOCPAELV 194

RESULT 11
US-09-057-088-10
Sequence 10, Application US/09057088
Patent No. 6114296
GENERAL INFORMATION:
APPLICANT: Schuelein, Martin
APPLICANT: Fredholm, Henrik
APPLICANT: Hjorth, Carsten
APPLICANT: Rasmussen, Grethe
APPLICANT: Nielsen, Egon
APPLICANT: Rosholm, Peter
TITLE OF INVENTION: Cellulase Variants
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6114296o No. 6114296disk of No. 6114296th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,088
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,777

```

; FILING DATE: 05-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GREGG, VALETA A
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 3913-504-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-868-9655
; TRLEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 284 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6114296e
; US-09-057-088-10

```

```

Query Match          64.8%; Score 35; DB 4; Length 284;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 QLPCPAELL 9
        1:|||||:
DB      186 QVQCPAEV 194

```

```

RESULT 12
US-09-230-225B-6
; Sequence 6, Application US/09230225B
; Patent No. 6403362
; GENERAL INFORMATION:
; APPLICANT: Meiji Seika Kaisha, Ltd.
; TITLE OF INVENTION: Systems for the Mass Production of Proteins or Peptides by Micro
; FILE OF INVENTION: of the Genus Humicola
; FILE REFERENCE: VX990054
; CURRENT APPLICATION NUMBER: US/09/230,225B
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Humicola insolens
; US-09-230-225B-6

```

```

Query Match          64.8%; Score 35; DB 4; Length 285;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 QLPCPAELL 9
        1:|||||:
DB      187 QVQCPAEV 195

```

```

RESULT 13
US-09-254-733-3
; Sequence 3, Application US/09254733
; Patent No. 6277596
; GENERAL INFORMATION:
; APPLICANT: MATANABE, MANABU
; APPLICANT: MORIYA, TATSUKI
; APPLICANT: AOYAGI, KAORU
; APPLICANT: SUMIDA, NAOMI
; APPLICANT: MURAKAMI, TAKESHI
; TITLE OF INVENTION: REGULATORY SEQUENCE OF CELLULOSE CBH1 GENES ORIGINATING
; TITLE OF INVENTION: IN TRICHODERMA VIRIDE AND SYSTEM FOR MASS-PRODUCING
; FILE REFERENCE: 99-0266*/LC(WMC)/00144
; CURRENT APPLICATION NUMBER: US/09/254,733
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 52

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Humicola insolens
; US-09-254-733-3

```

```

Query Match          64.8%; Score 35; DB 4; Length 286;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 QLPCPAELL 9
        1:|||||:
DB      188 QVQCPAEV 196

```

```

RESULT 14
US-09-189-060B-72
; Sequence 72, Application US/09189060B
; Patent No. 6270968
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Sandal, Thomas
; APPLICANT: Kauppinen, Markus
; APPLICANT: Borge, Diderichsen
; TITLE OF INVENTION: Method of Providing No. 6270968e1 DNA Sequences
; FILE REFERENCE: 4772.204-US
; CURRENT APPLICATION NUMBER: US/09/189,060B
; CURRENT FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/DK97/00216
; PRIOR FILING DATE: 1997-05-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Hybrid
; US-09-189-060B-72

```

```

Query Match          64.8%; Score 35; DB 4; Length 304;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 QLPCPAELL 9
        1:|||||:
DB      206 QVQCPAEV 214

```

```

RESULT 15
US-08-090-013-2
; Sequence 2, Application US/08090013
; Patent No. 5443750
; GENERAL INFORMATION:
; APPLICANT: CONVENTS, ANDRE C
; APPLICANT: BUSCH, ALFRED
; APPLICANT: BAECK, ANDRE C
; TITLE OF INVENTION: DETERGENT COMPOSITION WITH HIGH ACTIVITY
; TITLE OF INVENTION: CELLULOSE AND SOFTENING CLAYS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE PROCTER & GAMBLE COMPANY
; STREET: 5299 SPRING GROVE AVENUE
; CITY: CINCINNATI
; STATE: OHIO
; COUNTRY: USA
; ZIP: 45217
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/090,013

```


; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91202880.0
; FILING DATE: 06-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PATEL, KEN K.
; REGISTRATION NUMBER: 33,988
; REFERENCE/DOCKET NUMBER: CM393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-627-6437
; TELEFAX: 513-627-4854
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-090-013-2

Query Match 64.8%; Score 35; DB 1; Length 305;
Best Local Similarity 66.7%; Pred. NO. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 QLPCPAELI 9
1:|||||
Db 207 QVCCPAELV 215

Search completed: July 9, 2003, 12:03:34
Job time : 23.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:02:11 ; Search time 40.5 Seconds
(without alignments)
28.746 Million cell updates/sec

Title: US-09-851-058-1
Perfect score: 54
Sequence: 1 QPCPAELLR 10

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEM_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	75.9	78	10	US-09-925-297-752
2	41	75.9	180	10	US-09-925-301-1173
3	41	75.9	445	9	US-10-177-293-296
4	39	72.2	20	10	US-09-841-132-242
5	39	72.2	20	10	US-09-841-132-243
6	39	72.2	235	9	US-09-261-329-10
7	39	72.2	553	10	US-09-841-132-441
8	38	70.4	41	10	US-09-864-761-43975
9	37	68.5	87	10	US-09-864-761-6356
10	37	68.5	284	10	US-09-815-242-11778
11	37	68.5	420	10	US-09-946-142-2
12	36	66.7	106	9	US-09-809-391-341
13	35	64.8	100	9	US-09-764-891-2875
14	35	64.8	118	9	US-09-764-868-1062
15	35	64.8	118	9	US-09-955-999-71
16	35	64.8	123	9	US-09-764-868-630
17	35	64.8	128	9	US-09-764-868-1074
18	35	64.8	128	9	US-09-955-999-102
19	35	64.8	143	9	US-09-764-868-1075

20	35	64.8	143	9	US-09-955-999-103	Sequence 103, App
21	35	64.8	302	9	US-09-261-329-1	Sequence 1, Appl
22	35	64.8	305	9	US-10-138-870-2	Sequence 2, Appl
23	35	64.8	305	9	US-10-156-761-11748	Sequence 11748, A
24	35	64.8	305	10	US-09-735-787-2	Sequence 2, Appl
25	35	64.8	308	9	US-10-062-548-60	Sequence 60, Appl
26	35	64.8	344	9	US-10-146-772-268	Sequence 268, App
27	35	64.8	417	9	US-09-738-676-4642	Sequence 4642, Ap
28	35	64.8	2504	9	US-10-242-056-12	Sequence 12, Appl
29	35	64.8	2504	10	US-09-817-514-8	Sequence 8, Appl
30	35	63.9	151	9	US-09-187-693-58	Sequence 58, Appl
31	34	63.0	20	10	US-09-841-132-241	Sequence 241, App
32	34	63.0	36	10	US-09-860-232A-11	Sequence 11, Appl
33	34	63.0	65	10	US-09-864-761-35663	Sequence 35663, A
34	34	63.0	114	9	US-10-097-065-245	Sequence 245, App
35	34	63.0	176	9	US-09-974-879-399	Sequence 399, App
36	34	63.0	176	9	US-09-305-736-399	Sequence 399, App
37	34	63.0	318	9	US-10-156-761-9749	Sequence 9749, Ap
38	34	63.0	448	9	US-10-156-761-9046	Sequence 9046, Ap
39	34	63.0	506	9	US-10-288-222A-24	Sequence 24, Appl
40	34	63.0	506	10	US-09-776-705-5	Sequence 5, Appl
41	34	63.0	521	10	US-09-925-297-474	Sequence 474, App
42	34	63.0	556	10	US-09-841-132-398	Sequence 398, App
43	34	63.0	825	9	US-10-010-802-3	Sequence 3, Appl
44	34	63.0	825	10	US-09-785-934-2	Sequence 2, Appl
45	34	63.0	825	10	US-09-921-667-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-925-297-752
Sequence 752, Application US/09925297
Patent No. US20020081659A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA105
CURRENT APPLICATION NUMBER: US/09/925, 297
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05969
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124, 270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 928
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 752
LENGTH: 78
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (72)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-752

Query Match 75.9% Score 41: DB 10: Length 78:
Best Local Similarity 87.5% Pred. No. 5.9:
Matches 7: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 2 LQCPAELL 9
|||||||
DB 54 LQCPAELL 61

RESULT 2
US-09-925-301-1173
Sequence 1173, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1173
LENGTH: 180
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (171)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1173

Query Match 75.9%; Score 41; DB 10; Length 180;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 QPCPAELL 9
DB 24 LQCPAEFLV 31

RESULT 3
US-10-177-293-296
Sequence 296, Application US/10177293
Publication No. US2003024128A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Ghatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamathkar, Shubhangl
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobagyl, Gabriel N.
APPLICANT: Pusztai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 296
LENGTH: 445
TYPE: PRT

ORGANISM: Homo sapiens
US-10-177-293-296

Query Match 75.9%; Score 41; DB 9; Length 445;
Best Local Similarity 87.5%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPCPAELL 9
DB 289 LQCPAEFLV 296

RESULT 4
US-09-841-132-242
Sequence 242, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Skelky, Yasir A.W.
TITLE OF INVENTION: PROBST, PETER
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841,132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 242
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Made in a lab
US-09-841-132-242

Query Match 72.2%; Score 39; DB 10; Length 20;
Best Local Similarity 70.0%; Pred. No. 3.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QPCPAELLR 10
DB 7 QPCPAEFLVR 16

RESULT 5
US-09-841-132-243
Sequence 243, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Skelky, Yasir A.W.
TITLE OF INVENTION: PROBST, PETER
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841,132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 243
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Made in a lab
US-09-841-132-243

Query Match 72.2%; Score 39; DB 10; Length 20;
Best Local Similarity 70.0%; Pred. No. 3.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 QPCPAELLR 10
DB 7 QPCPAEFLVR 16

```
Db          3  OLPCAEFVR 12

RESULT 6
US-09-261-329-10
; Sequence 10, Application US/09261329
; Publication No. US2003092097A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Kim
; APPLICANT: Schulteis, Martin
; APPLICANT: Christensen, Lars
; APPLICANT: Damgaard, Bo
; APPLICANT: Von Der Osten, Claus
; TITLE OF INVENTION: Cellulase Variants
; FILE REFERENCE: 4887.204-US
; CURRENT APPLICATION NUMBER: US/09/261,329
; CURRENT FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 1013/96
; EARLIER FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Cellulase variants
US-09-261-329-10

Query Match
Best Local Similarity 72.2%; Score 39; DB 9; Length 235;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy          1  OLPCPAEL 8
Db          219 EVPCPAEL 226

RESULT 7
US-09-841-132-441
; Sequence 441, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhalla, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 441
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-441

Query Match
Best Local Similarity 72.2%; Score 39; DB 10; Length 553;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy          1  OLPCPAELLR 10
Db          145 OLPCAEFVR 154

RESULT 8
US-09-864-761-43975
; Sequence 43975, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 43975
LENGTH: 41
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005215.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92
OTHER INFORMATION: SWISSPROT HIT: P53233, EVALUATE 4.80e+00
US-09-864-761-43975

Query Match
Best Local Similarity 70.4%; Score 38; DB 10; Length 41;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy          2  LPCPAELLR 10
Db          22 LPCPDVLR 30

RESULT 9
US-09-864-761-46356
; Sequence 46356, Application US/09864761
; Patent No. US20020048763A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-x-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46356
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ACC007845.12
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EST_HUMAN HIT: AA63747.1, EVALUATE 1.00e-25
; OTHER INFORMATION: SWISSPROT HIT: Q61670, EVALUATE 5.00e-02
US-09-864-761-46356

Query Match
Best Local Similarity 68.5%; Score 37; DB 10; Length 87;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLPCPAE 7
:|||||
Db 21 RLPCPAE 27

RESULT 10
US-09-815-242-11778
; Sequence 11778, Application US/09815242
; Patent No. US20020061569A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11778
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11778

Query Match
Best Local Similarity 68.5%; Score 37; DB 10; Length 284;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRPCAELLR 10
:|||||
Db 211 MPVPAELLR 219

RESULT 11
US-09-946-142-2
; Sequence 2, Application US/09946142
; Patent No. US20020055152A1
; GENERAL INFORMATION:
; APPLICANT: FARMICK, Mike, et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE 11DD2 GENE
; FILE REFERENCE: 032301 WD 193
; CURRENT APPLICATION NUMBER: US/09/946,142
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-946-142-2

Query Match
Best Local Similarity 68.5%; Score 37; DB 10; Length 420;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLPCPAELL 9
:|||||
Db 5 QLPNPAELL 13

RESULT 12
```

US-09-809-391-341
; Sequence 341, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 341
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (102)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-809-391-341

Query Match

66.7%; Score 36; DB 9; Length 106;
Best Local Similarity 57.1%; Pred. No. 55;

Matches 8; Conservative 2; Mismatches 0; Indels 4; Gaps 1;

OY 1 QLPCLP---AELLR 10
| | | | | : | | | |

Db 36 QLPCLPLLSAQLLK 49

RESULT 13
US-09-764-891-2875
; Sequence 2875, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2875
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (52)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-2875

Query Match

64.8%; Score 35; DB 9; Length 100;
Best Local Similarity 85.7%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 PCPAELL 9
| | | | |

Db 90 PCPAHL 96

RESULT 14

US-09-764-868-1062
; Sequence 1062, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1062
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (101)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1062

Query Match

64.8%; Score 35; DB 9; Length 118;
Best Local Similarity 66.7%; Pred. No. 91;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCLPALL 9
| | | | |

Db 8 QLPCLPASVL 16

RESULT 15
US-09-955-999-71
; Sequence 71, Application US/09955999
; Publication No. US2003036505A1
; GENERAL INFORMATION:
; APPLICANT: Barash et al.
; TITLE OF INVENTION: Signal Transduction Pathway Component Polynucleotides, Polypep
; FILE REFERENCE: PT086P1
; CURRENT APPLICATION NUMBER: US/09/955,999
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (101)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-955-999-71

Query Match

64.8%; Score 35; DB 9; Length 118;
Best Local Similarity 66.7%; Pred. No. 91;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCLPALL 9
| | | | |

Db 8 QLPCLPASVL 16

Search completed: July 9, 2003, 12:17:58
Job time : 41.5 secs

This Page Blank (supro).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:00:11 ; Search time 286.5 Seconds

(Without alignments)
22.504 Million cell updates/sec

Title: US-09-851-058-1

Perfect score: 54

Sequence: 1 QLRCPAELLR 10

Scoring table: BLOSUM62

Searched: Gap0 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA Main:*

1: /cgn2_6/ptodata/1/paa/PCPUS.COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US06.COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US081.COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US082.COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US083.COMB.pep:*

8: /cgn2_6/ptodata/1/paa/US084.COMB.pep:*

9: /cgn2_6/ptodata/1/paa/US085.COMB.pep:*

10: /cgn2_6/ptodata/1/paa/US086.COMB.pep:*

11: /cgn2_6/ptodata/1/paa/US087.COMB.pep:*

12: /cgn2_6/ptodata/1/paa/US088.COMB.pep:*

13: /cgn2_6/ptodata/1/paa/US089.COMB.pep:*

14: /cgn2_6/ptodata/1/paa/US090.COMB.pep:*

15: /cgn2_6/ptodata/1/paa/US091.COMB.pep:*

16: /cgn2_6/ptodata/1/paa/US092.COMB.pep:*

17: /cgn2_6/ptodata/1/paa/US093.COMB.pep:*

18: /cgn2_6/ptodata/1/paa/US094.COMB.pep:*

19: /cgn2_6/ptodata/1/paa/US095.COMB.pep:*

20: /cgn2_6/ptodata/1/paa/US096.COMB.pep:*

21: /cgn2_6/ptodata/1/paa/US097.COMB.pep:*

22: /cgn2_6/ptodata/1/paa/US098.COMB.pep:*

23: /cgn2_6/ptodata/1/paa/US099.COMB.pep:*

24: /cgn2_6/ptodata/1/paa/US100.COMB.pep:*

25: /cgn2_6/ptodata/1/paa/US101.COMB.pep:*

26: /cgn2_6/ptodata/1/paa/US102.COMB.pep:*

27: /cgn2_6/ptodata/1/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	471	21	US-09-791-537-104983
2	54	100.0	471	21	US-09-791-537-143787
3	54	100.0	476	21	US-09-791-537-143801
4	54	100.0	476	22	US-09-801-563-43
5	42	77.8	328	1	PCT-US01-14827-10106
6	41	75.9	50	1	PCT-US01-01354-16078

7	41	75.9	50	21	US-09-764-905-16078	Sequence 16078, A
8 <td>41<th>75.9</th><td>50<td>24</td><td>US-10-092-399-16078</td><td>Sequence 16078, A</td></td></td>	41 <th>75.9</th> <td>50<td>24</td><td>US-10-092-399-16078</td><td>Sequence 16078, A</td></td>	75.9	50 <td>24</td> <td>US-10-092-399-16078</td> <td>Sequence 16078, A</td>	24	US-10-092-399-16078	Sequence 16078, A
9 <td>41<th>75.9</th><td>78</td><td>1</td><td>PCT-US00-05989-752</td><td>Sequence 752, App</td></td>	41 <th>75.9</th> <td>78</td> <td>1</td> <td>PCT-US00-05989-752</td> <td>Sequence 752, App</td>	75.9	78	1	PCT-US00-05989-752	Sequence 752, App
10 <td>41<th>75.9</th><td>78</td><td>23</td><td>US-09-925-297-752</td><td>Sequence 752, App</td></td>	41 <th>75.9</th> <td>78</td> <td>23</td> <td>US-09-925-297-752</td> <td>Sequence 752, App</td>	75.9	78	23	US-09-925-297-752	Sequence 752, App
11 <td>41<th>75.9</th><td>180</td><td>1</td><td>PCT-US00-05882-1173</td><td>Sequence 1173, App</td></td>	41 <th>75.9</th> <td>180</td> <td>1</td> <td>PCT-US00-05882-1173</td> <td>Sequence 1173, App</td>	75.9	180	1	PCT-US00-05882-1173	Sequence 1173, App
12 <td>41<th>75.9</th><td>180</td><td>23</td><td>US-09-925-301-1173</td><td>Sequence 1173, App</td></td>	41 <th>75.9</th> <td>180</td> <td>23</td> <td>US-09-925-301-1173</td> <td>Sequence 1173, App</td>	75.9	180	23	US-09-925-301-1173	Sequence 1173, App
13 <td>41<th>75.9</th><td>412</td><td>1</td><td>PCT-US02-31095-22</td><td>Sequence 22, App</td></td>	41 <th>75.9</th> <td>412</td> <td>1</td> <td>PCT-US02-31095-22</td> <td>Sequence 22, App</td>	75.9	412	1	PCT-US02-31095-22	Sequence 22, App
14 <td>41<th>75.9</th><td>445</td><td>1</td><td>PCT-US02-19668-296</td><td>Sequence 296, App</td></td>	41 <th>75.9</th> <td>445</td> <td>1</td> <td>PCT-US02-19668-296</td> <td>Sequence 296, App</td>	75.9	445	1	PCT-US02-19668-296	Sequence 296, App
15 <td>41<th>75.9</th><td>445</td><td>10</td><td>US-08-691-814-6</td><td>Sequence 6, App</td></td>	41 <th>75.9</th> <td>445</td> <td>10</td> <td>US-08-691-814-6</td> <td>Sequence 6, App</td>	75.9	445	10	US-08-691-814-6	Sequence 6, App
16 <td>41<th>75.9</th><td>445</td><td>10</td><td>US-08-691-814-6</td><td>Sequence 296, App</td></td>	41 <th>75.9</th> <td>445</td> <td>10</td> <td>US-08-691-814-6</td> <td>Sequence 296, App</td>	75.9	445	10	US-08-691-814-6	Sequence 296, App
17 <td>41<th>75.9</th><td>445</td><td>25</td><td>US-10-177-293-296</td><td>Sequence 296, App</td></td>	41 <th>75.9</th> <td>445</td> <td>25</td> <td>US-10-177-293-296</td> <td>Sequence 296, App</td>	75.9	445	25	US-10-177-293-296	Sequence 296, App
18 <td>41<th>75.9</th><td>468</td><td>21</td><td>US-09-791-537-24715</td><td>Sequence 24715, A</td></td>	41 <th>75.9</th> <td>468</td> <td>21</td> <td>US-09-791-537-24715</td> <td>Sequence 24715, A</td>	75.9	468	21	US-09-791-537-24715	Sequence 24715, A
19 <td>40<th>74.1</th><td>366</td><td>21</td><td>US-09-791-537-71757</td><td>Sequence 71757, A</td></td>	40 <th>74.1</th> <td>366</td> <td>21</td> <td>US-09-791-537-71757</td> <td>Sequence 71757, A</td>	74.1	366	21	US-09-791-537-71757	Sequence 71757, A
20 <td>39<th>72.2</th><td>20</td><td>18</td><td>US-09-454-684-242</td><td>Sequence 242, App</td></td>	39 <th>72.2</th> <td>20</td> <td>18</td> <td>US-09-454-684-242</td> <td>Sequence 242, App</td>	72.2	20	18	US-09-454-684-242	Sequence 242, App
21 <td>39<th>72.2</th><td>20</td><td>18</td><td>US-09-454-684-242</td><td>Sequence 243, App</td></td>	39 <th>72.2</th> <td>20</td> <td>18</td> <td>US-09-454-684-242</td> <td>Sequence 243, App</td>	72.2	20	18	US-09-454-684-242	Sequence 243, App
22 <td>39<th>72.2</th><td>20</td><td>18</td><td>US-09-454-684-242</td><td>Sequence 242, App</td></td>	39 <th>72.2</th> <td>20</td> <td>18</td> <td>US-09-454-684-242</td> <td>Sequence 242, App</td>	72.2	20	18	US-09-454-684-242	Sequence 242, App
23 <td>39<th>72.2</th><td>20</td><td>18</td><td>US-09-454-684-242</td><td>Sequence 243, App</td></td>	39 <th>72.2</th> <td>20</td> <td>18</td> <td>US-09-454-684-242</td> <td>Sequence 243, App</td>	72.2	20	18	US-09-454-684-242	Sequence 243, App
24 <td>39<th>72.2</th><td>20</td><td>19</td><td>US-09-598-419-242</td><td>Sequence 242, App</td></td>	39 <th>72.2</th> <td>20</td> <td>19</td> <td>US-09-598-419-242</td> <td>Sequence 242, App</td>	72.2	20	19	US-09-598-419-242	Sequence 242, App
25 <td>39<th>72.2</th><td>20</td><td>19</td><td>US-09-598-419-242</td><td>Sequence 243, App</td></td>	39 <th>72.2</th> <td>20</td> <td>19</td> <td>US-09-598-419-242</td> <td>Sequence 243, App</td>	72.2	20	19	US-09-598-419-242	Sequence 243, App
26 <td>39<th>72.2</th><td>20</td><td>22</td><td>US-09-841-132-242</td><td>Sequence 242, App</td></td>	39 <th>72.2</th> <td>20</td> <td>22</td> <td>US-09-841-132-242</td> <td>Sequence 242, App</td>	72.2	20	22	US-09-841-132-242	Sequence 242, App
27 <td>39<th>72.2</th><td>20</td><td>22</td><td>US-09-841-132-242</td><td>Sequence 243, App</td></td>	39 <th>72.2</th> <td>20</td> <td>22</td> <td>US-09-841-132-242</td> <td>Sequence 243, App</td>	72.2	20	22	US-09-841-132-242	Sequence 243, App
28 <td>39<th>72.2</th><td>51</td><td>13</td><td>US-08-905-279-359</td><td>Sequence 359, App</td></td>	39 <th>72.2</th> <td>51</td> <td>13</td> <td>US-08-905-279-359</td> <td>Sequence 359, App</td>	72.2	51	13	US-08-905-279-359	Sequence 359, App
29 <td>39<th>72.2</th><td>51</td><td>19</td><td>US-09-547-599C-2758</td><td>Sequence 2758, App</td></td>	39 <th>72.2</th> <td>51</td> <td>19</td> <td>US-09-547-599C-2758</td> <td>Sequence 2758, App</td>	72.2	51	19	US-09-547-599C-2758	Sequence 2758, App
30 <td>39<th>72.2</th><td>132</td><td>1</td><td>PCT-US00-30674-115</td><td>Sequence 115, App</td></td>	39 <th>72.2</th> <td>132</td> <td>1</td> <td>PCT-US00-30674-115</td> <td>Sequence 115, App</td>	72.2	132	1	PCT-US00-30674-115	Sequence 115, App
31 <td>39<th>72.2</th><td>132</td><td>1</td><td>PCT-US01-11988-1180</td><td>Sequence 1180, App</td></td>	39 <th>72.2</th> <td>132</td> <td>1</td> <td>PCT-US01-11988-1180</td> <td>Sequence 1180, App</td>	72.2	132	1	PCT-US01-11988-1180	Sequence 1180, App
32 <td>39<th>72.2</th><td>132</td><td>22</td><td>US-09-833-245-1180</td><td>Sequence 1180, App</td></td>	39 <th>72.2</th> <td>132</td> <td>22</td> <td>US-09-833-245-1180</td> <td>Sequence 1180, App</td>	72.2	132	22	US-09-833-245-1180	Sequence 1180, App
33 <td>39<th>72.2</th><td>133</td><td>1</td><td>PCT-US00-30674-82</td><td>Sequence 82, App</td></td>	39 <th>72.2</th> <td>133</td> <td>1</td> <td>PCT-US00-30674-82</td> <td>Sequence 82, App</td>	72.2	133	1	PCT-US00-30674-82	Sequence 82, App
34 <td>39<th>72.2</th><td>133</td><td>1</td><td>PCT-US01-11988-1183</td><td>Sequence 1183, App</td></td>	39 <th>72.2</th> <td>133</td> <td>1</td> <td>PCT-US01-11988-1183</td> <td>Sequence 1183, App</td>	72.2	133	1	PCT-US01-11988-1183	Sequence 1183, App
35 <td>39<th>72.2</th><td>133</td><td>22</td><td>US-09-833-245-1183</td><td>Sequence 1183, App</td></td>	39 <th>72.2</th> <td>133</td> <td>22</td> <td>US-09-833-245-1183</td> <td>Sequence 1183, App</td>	72.2	133	22	US-09-833-245-1183	Sequence 1183, App
36 <td>39<th>72.2</th><td>235</td><td>16</td><td>US-09-261-329-10</td><td>Sequence 10, App</td></td>	39 <th>72.2</th> <td>235</td> <td>16</td> <td>US-09-261-329-10</td> <td>Sequence 10, App</td>	72.2	235	16	US-09-261-329-10	Sequence 10, App
37 <td>39<th>72.2</th><td>553</td><td>22</td><td>US-09-841-132-441</td><td>Sequence 441, App</td></td>	39 <th>72.2</th> <td>553</td> <td>22</td> <td>US-09-841-132-441</td> <td>Sequence 441, App</td>	72.2	553	22	US-09-841-132-441	Sequence 441, App
38 <td>39<th>72.2</th><td>557</td><td>16</td><td>US-09-201-1228A-1035</td><td>Sequence 1035, App</td></td>	39 <th>72.2</th> <td>557</td> <td>16</td> <td>US-09-201-1228A-1035</td> <td>Sequence 1035, App</td>	72.2	557	16	US-09-201-1228A-1035	Sequence 1035, App
39 <td>38<th>70.4</th><td>41</td><td>1</td><td>PCT-US01-00663-33358</td><td>Sequence 33358, A</td></td>	38 <th>70.4</th> <td>41</td> <td>1</td> <td>PCT-US01-00663-33358</td> <td>Sequence 33358, A</td>	70.4	41	1	PCT-US01-00663-33358	Sequence 33358, A
40 <td>38<th>70.4</th><td>41</td><td>22</td><td>US-09-864-761-43375</td><td>Sequence 43375, A</td></td>	38 <th>70.4</th> <td>41</td> <td>22</td> <td>US-09-864-761-43375</td> <td>Sequence 43375, A</td>	70.4	41	22	US-09-864-761-43375	Sequence 43375, A
41 <td>38<th>70.4</th><td>41</td><td>25</td><td>US-10-182-993-32338</td><td>Sequence 32338, A</td></td>	38 <th>70.4</th> <td>41</td> <td>25</td> <td>US-10-182-993-32338</td> <td>Sequence 32338, A</td>	70.4	41	25	US-10-182-993-32338	Sequence 32338, A
42 <td>38<th>70.4</th><td>41</td><td>26</td><td>US-10-203-134-33164</td><td>Sequence 33164, A</td></td>	38 <th>70.4</th> <td>41</td> <td>26</td> <td>US-10-203-134-33164</td> <td>Sequence 33164, A</td>	70.4	41	26	US-10-203-134-33164	Sequence 33164, A
43 <td>38<th>70.4</th><td>41</td><td>26</td><td>US-10-203-135-32353</td><td>Sequence 32353, A</td></td>	38 <th>70.4</th> <td>41</td> <td>26</td> <td>US-10-203-135-32353</td> <td>Sequence 32353, A</td>	70.4	41	26	US-10-203-135-32353	Sequence 32353, A
44 <td>38<th>70.4</th><td>41</td><td>26</td><td>US-10-203-136-33211</td><td>Sequence 33211, A</td></td>	38 <th>70.4</th> <td>41</td> <td>26</td> <td>US-10-203-136-33211</td> <td>Sequence 33211, A</td>	70.4	41	26	US-10-203-136-33211	Sequence 33211, A
45 <td>38<th>70.4</th><td>41</td><td>26</td><td>US-10-203-137-33358</td><td>Sequence 33358, A</td></td>	38 <th>70.4</th> <td>41</td> <td>26</td> <td>US-10-203-137-33358</td> <td>Sequence 33358, A</td>	70.4	41	26	US-10-203-137-33358	Sequence 33358, A

ALIGNMENTS

RESULT 1

US-09-791-537-104983

Sequence 104983, Application US/09/91537

GENERAL INFORMATION:

APPLICANT: Biomomix, Inc.

APPLICANT: Danzer, Joseph

APPLICANT: Debe, Derek

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 104983

LENGTH: 471

TYPE: PRT

ORGANISM: Escherichia coli

US-09-791-537-104983

Query Match 100.0%; Score 54; DB 21; Length 471;

Best Local Similarity 100.0%; Pred. No. 4.6;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLRCPAELLR 10

Db 410 QLRCPAELLR 419

RESULT 2
US-09-791-537-143787
; Sequence 143787, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 143787
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-791-537-143787

Query Match
Best Local Similarity 100.0%; Score 54; DB 21; Length 471;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCPAELLR 10
DB 410 QLPCPAELLR 419

RESULT 3
US-09-791-537-143801
; Sequence 143801, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 143801
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-791-537-143801

Query Match
Best Local Similarity 100.0%; Score 54; DB 21; Length 476;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCPAELLR 10
DB 415 QLPCPAELLR 424

RESULT 4
US-09-801-563-43
; Sequence 43, Application US/09801563
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al.
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PR2-043
; CURRENT APPLICATION NUMBER: US/09/801,563
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 476

; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-801-563-43

Query Match
Best Local Similarity 100.0%; Score 54; DB 22; Length 476;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCPAELLR 10
DB 415 QLPCPAELLR 424

RESULT 5
PCT-US01-14827-10106
; Sequence 10106, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 10106
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (265)-(283)
; OTHER INFORMATION: Zinc finger, CCHC class domain identified by Pfam, accession PCT-US01-14827-10106

Query Match
Best Local Similarity 77.8%; Score 42; DB 1; Length 328;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 QLPCPAELLR 10
DB 45 KLDCPAELLR 54

RESULT 6
PCT-US01-01354-16078
; Sequence 16078, Application PC/TUS0101354
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC004PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01354
; CURRENT FILING DATE: 2001-03-17
; NUMBER OF SEQ ID NOS: 42506
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16078
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01354-16078

Query Match
Best Local Similarity 75.9%; Score 41; DB 1; Length 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCPAE 7
IIIIIIII

Db 14 QLCPPAE 20

RESULT 7
US-09-764-905-16078
Sequence 16078, Application US/09764905
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antipodles
FILE REFERENCE: PC004
CURRENT APPLICATION NUMBER: US/09/764,905
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368

PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17

;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,065
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,398

Query Match 75.9% Score 41; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCPAE 7
Db 14 QLPCPAE 20

RESULT 8
US-10-092-399-16078
; Sequence 16078, Application US/100923399
; GENERAL INFORMATION:

;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC004C1
;; CURRENT APPLICATION NUMBER: US/10/092,399
;; CURRENT FILING DATE: 2002-03-07
;; NUMBER OF SEQ ID NOS: 42506
;; PRIOR APPLICATION removed - See File Wrapper or Palm
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 16078
;; LENGTH: 50
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc-feature
;; LOCATION: (50)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-399-16078

Query Match 75.9% Score 41; DB 24; Length 50;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCPAE 7
Db 14 QLPCPAE 20

RESULT 9
PCT-US00-05989-752
; Sequence 752, Application PC/TUS0005989
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Pancreas and Pancreatic Cancer Associated Gene Sequences
; FILE REFERENCE: PA105PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05989
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 752
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05989-752

Query Match 75.9% Score 41; DB 1; Length 78;
Best Local Similarity 87.5%; Pred. No. 1,1e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAEL 9
Db 54 LPCPAELV 61

RESULT 10
US-09-925-297-752
; Sequence 752, Application US/09925297
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

```
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 752
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (172)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-752

Query Match          75.9%; Score 41; DB 23; Length 78;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 LPCPAELL 9
Db      54 LPCPAELV 61

RESULT 11
PCT-US00-05882-1173
; Sequence 1173, Application PC/TUS0005882
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: PA106PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05882
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 1173
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (171)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05882-1173

Query Match          75.9%; Score 41; DB 1; Length 180;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 LPCPAELL 9
Db      24 LPCPAELV 31

RESULT 12
US-09-925-301-1173
; Sequence 1173, Application US/09925301
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 1173
; LENGTH: 180
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (171)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1173

Query Match          75.9%; Score 41; DB 23; Length 180;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 LPCPAELL 9
Db      24 LPCPAELV 31

RESULT 13
PCT-US02-31095-22
; Sequence 22, Application PC/TUS0231095
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yaida
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: BOROMSKY, Mark L.
; APPLICANT: CHAMLA, Narinder K.
; APPLICANT: ELIOTT, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: HAPALIA, April J.A.
; APPLICANT: ISON, Craig H.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KALAFUS, Daniel P.
; APPLICANT: LEHR-MASON, Patricia M.
; APPLICANT: LU, Dyung Alina M.
; APPLICANT: MAROUI, Joseph P.
; APPLICANT: NGUYEN, Danielle B.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: SAPPERSSTEIN, Stephanie K.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: TRAN,uyen K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuning
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Huibin
; APPLICANT: YUE, Henry
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEA
; FILE REFERENCE: PF-1213 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/31095
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/326,389
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/327,380
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/328,186
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/329,690
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/345,384
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/348,165
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/350,219
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/344,518
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/345,143
; PRIOR FILING DATE: 2001-11-09
```

PRIOR APPLICATION NUMBER: US 60/332,375
PRIOR FILING DATE: 2001-11-16
Remainder Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PERL Program
SEQ ID NO 22
LENGTH: 412
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 7503812CD1
PCT-US02-31095-22

Query Match 75.9%: Score 41; DB 1; Length 412;
Best Local Similarity 87.5%: Pred. No. 5.6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAELL 9
Db 256 LPCPAELV 263

RESULT 14
PCT-US02-19669-296
Sequence 296, Application PC/TUS0219669
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc. et al.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038PC
CURRENT APPLICATION NUMBER: PCT/US02/19669
PRIOR FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 296
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-19669-296

Query Match 75.9%: Score 41; DB 1; Length 445;
Best Local Similarity 87.5%: Pred. No. 5.6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAELL 9
Db 289 LPCPAELV 296

RESULT 15
US-08-691-814-6
Sequence 6, Application US/08691814
GENERAL INFORMATION:
APPLICANT: Rio, Marie-Christine
APPLICANT: Tomasello, Catherine
APPLICANT: Bassett, Paul
APPLICANT: Byrne, Jennifer
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful
TITLE OF INVENTION: as Leukemia Markers and in Breast Cancer Prognosis
NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
STREET: 1100 New York Ave, NW, Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/691,814
FILING DATE: 31-Jul-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,183
FILING DATE: 09-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1383, 0090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2543
TELEFAX: 202-371-2543
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-691-814-6

Query Match 75.9%: Score 41; DB 10; Length 445;
Best Local Similarity 87.5%: Pred. No. 5.6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAELL 9
Db 289 LPCPAELV 296

Search completed: July 9, 2003, 12:13:15
Job time : 289.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:00:51 ; Search time 94.5 Seconds
(without alignments)
27.740 Million cell updates/sec

Title: US-09-851-058-1
Perfect score: 54
Sequence: 1 QLPCPAELLR 10

Scoring table:
BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 1221054 seqs, 262139423 residues

Total number of hits satisfying chosen parameters: 1221054

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	10	2 PCT-US02-14369-1	Sequence 1, Appl
2	54	100.0	10	10 US-09-851-058-1	Sequence 12, Appl
3	54	100.0	471	12 US-10-254-763-12	Sequence 20678, A
4	45	83.3	79	2 PCT-US02-32727-20678	Sequence 20678, A
5	45	83.3	79	10 US-09-978-825-20678	Sequence 20678, A
6	45	83.3	79	12 US-10-057-498-20678	Sequence 20678, A
7	42	77.8	118	12 US-10-437-963-127901	Sequence 103514, A
8	42	77.8	431	12 US-10-437-963-103514	Sequence 57238, A
9	41	75.9	176	10 US-09-724-676-57238	Sequence 57238, A
10	41	75.9	176	10 US-09-724-676-57238	Sequence 57238, A
11	41	75.9	186	10 US-09-724-676-57248	Sequence 57248, A
12	41	75.9	186	10 US-09-724-676-57248	Sequence 57248, A
13	41	75.9	223	10 US-09-724-676-57240	Sequence 57240, A
14	41	75.9	223	10 US-09-724-676-57240	Sequence 57240, A
15	41	75.9	233	10 US-09-724-676-57239	Sequence 57239, A
16	41	75.9	233	10 US-09-724-676-57239	Sequence 57239, A
17	41	75.9	435	10 US-09-724-676-57242	Sequence 57242, A
18	41	75.9	435	10 US-09-724-676-57242	Sequence 57242, A
19	41	75.9	445	2 PCT-US02-19669A-296	Sequence 296, App

20	41	75.9	445	14 US-60-453-135-13494	Sequence 13494, A
21	41	75.9	445	14 US-60-453-050-13494	Sequence 13494, A
22	41	75.9	445	14 US-60-466-412-13494	Sequence 13494, A
23	41	75.9	475	10 US-09-724-676-57246	Sequence 57246, A
24	41	75.9	475	10 US-09-724-676-57246	Sequence 57246, A
25	41	75.9	485	10 US-09-724-676-57245	Sequence 57245, A
26	41	75.9	485	10 US-09-724-676-57245	Sequence 57245, A
27	41	75.9	560	10 US-09-724-676-57244	Sequence 57244, A
28	41	75.9	560	10 US-09-724-676-57244	Sequence 57244, A
29	41	75.9	570	10 US-09-724-676-57243	Sequence 57243, A
30	41	75.9	570	10 US-09-724-676-57243	Sequence 57243, A
31	41	75.9	600	10 US-09-724-676-57247	Sequence 57247, A
32	41	75.9	600	10 US-09-724-676-57247	Sequence 57247, A
33	41	75.9	610	10 US-09-724-676-57241	Sequence 57241, A
34	41	75.9	610	10 US-09-724-676-57241	Sequence 57241, A
35	41	75.9	1959	14 US-60-419-463-34	Sequence 34, Appl
36	40	74.1	352	12 US-10-424-599-245651	Sequence 245651, A
37	40	74.1	5149	12 US-10-282-122A-43633	Sequence 43633, A
38	39	72.2	52	12 US-10-424-599-254154	Sequence 254154, A
39	39	72.2	252	8 US-08-411-777A-8	Sequence 8, Appl
40	39	72.2	252	10 US-09-057-088A-8	Sequence 8, Appl
41	39	72.2	554	12 US-10-369-435-12	Sequence 12, Appl
42	38	70.4	122	12 US-10-424-599-153776	Sequence 153776, A
43	38	70.4	183	12 US-10-424-599-270126	Sequence 270126, A
44	38	70.4	211	12 US-10-425-114-49901	Sequence 49901, A
45	38	70.4	265	12 US-10-282-122A-55249	Sequence 55249, A

ALIGNMENTS

RESULT 1
PCT-US02-14369-1
Sequence 1, Application PC/TUS0214369

GENERAL INFORMATION:

APPLICANT: Parker, Kenneth

APPLICANT: Nadler, Timothy

APPLICANT: Vella, George

APPLICANT: Huang, Yulin

APPLICANT: Abersold, Rudolf

APPLICANT: Smolka, Marcus

TITLE OF INVENTION: Process for Analyzing Protein Samples

FILE REFERENCE: SYP-172

CURRENT APPLICATION NUMBER: PCT/US02/14369

CURRENT FILING DATE: 2003-02-05

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 1

LENGTH: 10

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: synthetic peptides

PCT-US02-14369-1

Query Match

Best Local Similarity 100.0%; Score 54; DB 2; Length 10;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLPCPAELLR 10

Db 1 QLPCPAELLR 10

RESULT 2

US-09-851-058-1

Sequence 1, Application US/09851058

GENERAL INFORMATION:

APPLICANT: Parker, Kenneth

APPLICANT: Nadler, Timothy

APPLICANT: Vella, George

APPLICANT: Huang, Yulin

APPLICANT: Abersold, Rudolf

APPLICANT: Smolka, Marcus
TITLE OF INVENTION: Process for Analyzing Protein Samples
FILE REFERENCE: SYP-172
CURRENT APPLICATION NUMBER: US/09/851,058
CURRENT FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptides
US-09-851-058-1

Query Match 100.0%; Score 54; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OLPCPAELLR 10
DB 1 OLPCPAELLR 10

RESULT 3
US-10-254-763-12
Sequence 12, Application US/10254763
GENERAL INFORMATION:
APPLICANT: TAKAGI, HIROSHI
APPLICANT: WADA, MASARU
APPLICANT: NAKAMORI, SHIGERU
TITLE OF INVENTION: L-Cysteine-Producing Bacterium and Method for Producing L-Cysteine
FILE REFERENCE: 228254050
CURRENT APPLICATION NUMBER: US/10/254,763
CURRENT FILING DATE: 2003-01-06
PRIOR APPLICATION NUMBER: JP 2001-302008
PRIOR FILING DATE: 2001-09-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 471
TYPE: PRT
ORGANISM: Escherichia coli
US-10-254-763-12

Query Match 100.0%; Score 54; DB 12; Length 471;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OLPCPAELLR 10
DB 410 OLPCPAELLR 419

RESULT 4
PCT-US02-32727-20678
Sequence 20678, Application PC/TUS0232727
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skelky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yanni
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shylan
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darrick
APPLICANT: Barth, Brenda
APPLICANT: Douglas, John
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: PCT/US02/32727
CURRENT FILING DATE: 2002-10-11
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 20678
LENGTH: 79
TYPE: PRT
ORGANISM: Propionibacterium
PCT-US02-32727-20678

Query Match 83.3%; Score 45; DB 2; Length 79;
Best Local Similarity 88.9%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPCPAELLR 10
DB 54 LPCPAELLR 62

RESULT 5
US-09-978-825-20678
Sequence 20678, Application US/09978825
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skelky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yanni
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shylan
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darrick
APPLICANT: Barth, Brenda
APPLICANT: Douglas, John
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: US/09/978,825
CURRENT FILING DATE: 2003-01-29
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 20678
LENGTH: 79
TYPE: PRT
ORGANISM: Propionibacterium
US-09-978-825-20678

Query Match 83.3%; Score 45; DB 10; Length 79;
Best Local Similarity 88.9%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPCPAELLR 10
DB 54 LPCPAELLR 62

RESULT 6
US-10-057-498-20678
Sequence 20678, Application US/10057498
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skelky, Yasir
APPLICANT: Persing, David
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514
CURRENT APPLICATION NUMBER: US/10/057,498
CURRENT FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 29212
SEQ ID NO 20678
LENGTH: 79
TYPE: PRT
ORGANISM: Propionibacterium

US-10-057-498-20678

Query Match 83.3%; Score 45; DB 12; Length 79;
Best Local Similarity 88.9%; Pred. No. 4.6;
Matches 8: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPCPAELLR 10
|||||
DB 54 LPCPTELLR 62

RESULT 7

US-10-437-963-127901

; Sequence 127901, Application US/10437963

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 127901

; LENGTH: 118

; TYPE: PRT

; ORGANISM: Oryza sativa

; OTHER INFORMATION: Clone ID: PAT_MRT4530_30306C.1.pep

US-10-437-963-127901

Query Match 77.8%; Score 42; DB 12; Length 118;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPCPAELLR 10
|||||
DB 14 LPCPARILR 22

RESULT 8

US-10-437-963-103514

; Sequence 103514, Application US/10437963

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 103514

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Oryza sativa

; NAME/KEY: unsure

; LOCATION: (1)..(431)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_100936C.1.pep

US-10-437-963-103514

Query Match 77.8%; Score 42; DB 12; Length 431;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 8: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QLPCPAELLR 10
|||||
DB 32 QLPCPLELNR 41

RESULT 9

US-09-724-676-57238

; Sequence 57238, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; SOFTWARE: PatentIn version 3.2

; NUMBER OF SEQ ID NOS: 97222

; SEQ ID NO 57238

; LENGTH: 176

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-57238

Query Match 75.9%; Score 41; DB 10; Length 176;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 7: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPCPAELL 9
|||||
DB 30 LPCPAELV 37

RESULT 10

US-09-724-676A-57238

; Sequence 57238, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 57238

; LENGTH: 176

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-57238

Query Match 75.9%; Score 41; DB 10; Length 176;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 7: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPCPAELL 9
|||||
DB 30 LPCPAELV 37

RESULT 11

US-09-724-676-57248

; Sequence 57248, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

```
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57248
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-57248
```

```
Query Match      75.9%; Score 41; DB 10; Length 186;
Best Local Similarity 87.5%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 LPCPAELL 9
Db      30 LPCPAELV 37
```

```
RESULT 12
US-09-724-676A-57248
; Sequence 57248, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57248
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-57248
```

```
Query Match      75.9%; Score 41; DB 10; Length 186;
Best Local Similarity 87.5%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 LPCPAELL 9
Db      30 LPCPAELV 37
```

```
RESULT 13
US-09-724-676-57240
; Sequence 57240, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57240
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-57240
```

```
Query Match      75.9%; Score 41; DB 10; Length 223;
Best Local Similarity 87.5%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 LPCPAELL 9
Db      77 LPCPAELV 84
```

```
RESULT 14
US-09-724-676A-57240
; Sequence 57240, Application US/09724676A
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57240
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-57240
```

```
Query Match      75.9%; Score 41; DB 10; Length 223;
Best Local Similarity 87.5%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 LPCPAELL 9
Db      77 LPCPAELV 84
```

```
RESULT 15
US-09-724-676-57239
; Sequence 57239, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57239
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-57239
```

```
Query Match      75.9%; Score 41; DB 10; Length 233;
Best Local Similarity 87.5%; Pred. No. 58;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 LPCPAELL 9
Db      77 LPCPAELV 84
```

```
Search completed: July 9, 2003, 12:16:30
Job time : 95.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: July 9, 2003, 11:58:06 ; Search time 18.5 Seconds
(without alignments)
51.965 Million cell updates/sec

Title: US-09-851-058-1
Sequence: 1 QLPCPALLR 10

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	476	1 WZEC	tryptophanase (EC
2	54	100.0	476	2 E91209	tryptophanase [imp
3	54	100.0	476	3 H86055	tryptophanase [imp
4	47	87.0	472	2 H82492	tryptophanase VCA0
5	41	75.9	445	2 I38027	MLN 64 protein - h
6	41	75.9	445	2 S47447	hypothetical prote
7	40	74.1	5149	2 H83345	probable non-ribos
8	39	72.2	234	2 H82730	ribonuclease HII X
9	39	72.2	511	2 S10527	endoglucanase B pr
10	39	72.2	547	1 A32244	60K cysteine-rich
11	39	72.2	547	2 B43584	60K cysteine-rich
12	39	72.2	547	2 D71515	60K cysteine-rich
13	38	70.5	642	2 E87644	sensory box histid
14	37	68.4	284	2 F83615	hypothetical prote
15	36	66.7	63	2 G84331	hypothetical prote
16	36	66.7	335	2 T25054	hypothetical prote
17	36	66.7	339	2 S62515	precortin-gly-depen
18	36	66.7	415	2 E83281	probable retroelem
19	36	66.7	976	2 C84500	hypothetical prote
20	35	64.8	118	2 S72922	hypothetical prote
21	35	64.8	146	2 T48034	hZIP transcription
22	35	64.8	281	2 B95279	conserved hypotet
23	35	64.8	590	2 T49672	related to a-arglin
24	35	64.8	832	2 S76815	hypothetical prote
25	34	63.0	141	2 T33212	hypothetical prote
26	34	63.0	196	2 A40623	heat shock protein
27	34	63.0	206	2 D90630	hypothetical prote
28	34	63.0	206	2 D85481	hypothetical prote
29	34	63.0	222	2 A59263	tetranspan TSPAN-2

ALIGNMENTS

RESULT 1

WZEC
tryptophanase (EC 4.1.99.1) - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 14-Nov-1983 #sequence, revision 05-Dec-1997 #text, change 01-Mar-2002

C:Accession: E65173; A91789; I41097; I69358; A01136

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.; Mau, B.; Shaoh, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: E65173

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-476 <BLAT>

A:Cross-References: GB:A000448; GB:000096; NID:g1790142; PIDN:AACT6731.1; PID:g17901

A:Experimental source: strain K-12, substrain MG1655

R:Deeley, M.C.; Yanofsky, C.

J. Bacteriol. 147, 787-796, 1981

A:Title: Nucleotide sequence of the structural gene for tryptophanase of Escherichia

A:Reference number: A91789; MUID:82007678; PMID:6268608

A:Accession: A91789

A:Molecule type: DNA

A:Residues: 6-141, 'T', '143', 'OG', '146-383', 'TG', '386-476 <DE>

A:Experimental source: strain K-12

A>Note: the authors translated the codons GAT, ACG, and CAG for residues 142, 144, an

R:Kagamiyama, H.; Matsubara, H.; Spell, E.E.

J. Biol. Chem. 247, 1576-1586, 1972

A:Title: The chemical structure of tryptophanase from Escherichia coli. III. Isolatio

A:Reference number: A92100; MUID:72134434; PMID:4551944

A:Contents: annotation; sequences of tryptic peptides; strain K-12

A>Note: Lys-275 binds pyridoxal 5'-phosphate

R:Tokushige, M.; Tsujimoto, N.; Oda, T.; Honda, T.; Yumoto, N.; Ito, S.; Yamamoto, M

Biochimie 71, 711-720, 1989

A:Title: Role of cysteine residues in tryptophanase for monovalent cation-induced act

A:Reference number: I41097; MUID:89323226; PMID:2502187

A:Accession: I41097

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 6-383, 'TG', '386-476 <RES>

A:Cross-References: EMBL:X15974; NID:g41935; PIDN:CAA34096.1; PID:g41936

R:Stewart, V.J.; Yanofsky, C.

J. Bacteriol. 164, 731-740, 1985

A:Title: Evidence for transcription antitermination control of tryptophanase operon e

A:Reference number: I54862; MUID:86033634; PMID:3902796

A:Accession: I54862

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 6-27 <RES>

A:Cross-References: GB:M11990; NID:g147999; PIDN:AAA24679.1; PID:g148001

C:Genetics:

A:Gene: tnaA

A:Map position: 83 min
C:Complex: homotetramer
C:Function:
A:Description: catalyzes the degradation of tryptophan to indole, pyruvate, and ammonia;
C:Superfamily: tryptophanase
C:Keyword: carbon-carbon lyase; homotetramer

Query Match 100.0%; Score 54; DB 1; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 415 QLPCPABELL 424

RESULT 2
E91209
tryptophanase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509952)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: E91209
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gawawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E91209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <STO>
A:Cross-references: GB:BA000007; PIDN:BAJ38068.1; PID:q13364120; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC4645
C:Superfamily: tryptophanase

Query Match 100.0%; Score 54; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 415 QLPCPABELL 424

RESULT 3
H86055
tryptophanase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H86055
R:Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Iller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoustis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A89480; MUID:21074935; PMID:11206551
A:Accession: H86055
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <STO>
A:Cross-references: GB:AE005174; NID:912518553; PIDN:AAG58908.1; GSPDB:GN00145; UMG:252
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: tnaA
C:Superfamily: tryptophanase

Query Match 100.0%; Score 54; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 415 QLPCPABELL 424

DB 415 QLPCPABELL 424

RESULT 4
B82492
tryptophanase VCA0161 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82492
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82492
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <HEI>
A:Cross-references: GB:AE004357; GB:AE003853; NID:96557547; PIDN:AAF96074.1; GSPDB:GN
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0161
A:Map position: 2
C:Superfamily: tryptophanase

Query Match 87.0%; Score 47; DB 2; Length 472;
Best Local Similarity 90.0%; Pred. No. 0.96;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 410 QHPCPABELL 419

RESULT 5
I38027
MLN 64 protein - human
C:Species: Homo sapiens (man)
C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 05-Nov-1999
C:Accession: I38027; S60862
R:Tomasetto, C.; Regnier, C.H.; Moog-Lutz, C.; Mattei, M.G.; Chenard, M.P.; Lidereau,
Genomics 28, 367-376, 1995
A>Title: Identification of four novel human genes amplified and overexpressed in brea
A:Reference number: 137080; MUID:96039245; PMID:7490069
A:Accession: I38027
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-445 <RES>
A:Cross-references: EMBL:X80198; NID:951278; PIDN:CAA56489.1; PID:951279
A:Note: submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Gene: MLN64

Query Match 75.9%; Score 41; DB 2; Length 445;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 289 LKCPABELL 296

RESULT 6
S47447
hypothetical protein YML102w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 26-May-2000
C:Accession: S47447
R:Barrell, B.G.
submitted to the EMBL Data Library, August 1994
A:Reference number: S47445
A:Accession: S47447

A:Molecule type: DNA
A:Residues: 1-468 <BAR>
A:Cross-references: EMBL:X80835; NID:9530339; PIDN:CAA56795.1; PID:9530342; MIPS:YML102w
C:Genetics:
A:Gene: SGD:CAAC2
A:Cross-references: SGD:S0004570; MIPS:YML102w
A:Map position: 13L
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match
Best Local Similarity 75.9%; Score 41; DB 2; Length 468;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 QLPCPAELR 10
Db 237 ELPCGPVLR 246

RESULT 7
F83345
probable non-ribosomal peptide synthetase PA2402 [Imported] - Pseudomonas aeruginosa (str
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83345
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Llm
; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83345
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5149 <STO>
A:Cross-references: GB:A0004667; GB:A0004091; NID:99948444; PIDN:AA05790.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2402
C:Superfamily: acyl carrier protein homology; acetate-CoA ligase homology
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:557-998/Domain: acetate-CoA ligase homology <ACLI>
F:1015-1082/Domain: acyl carrier protein homology <ACPI>
F:2034-2479/Domain: acetate-CoA ligase homology <ACPI>
F:2496-3564/Domain: acyl carrier protein homology <ACPI>
F:3004-3525/Domain: acetate-CoA ligase homology <ACPI>
F:3541-3608/Domain: acyl carrier protein homology <ACPI>
F:4583-5040/Domain: acetate-CoA ligase homology <ACPI>
F:5057-5125/Domain: acyl carrier protein homology <ACPI>
F:1047,2528,3573,5089/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match
Best Local Similarity 74.1%; Score 40; DB 2; Length 5149;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 QLPCPAELR 10
Db 2098 RLPCPAEYR 2107

RESULT 8
H82730
ribonuclease H1 XF1041 [Imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
C:Accession: H82730
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82730
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-234 <SIM>
A:Cross-references: GB:A0003941; GB:A0003849; NID:99105978; PIDN:AAF83851.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinolan, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; France, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C
; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa, V.E.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1041
C:Superfamily: ribonuclease H1

Query Match
Best Local Similarity 72.2%; Score 39; DB 2; Length 234;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 LRPCPAEL 9
Db 123 LRPCPAEL 130

RESULT 9
S10527
endoglucanase B precursor - Pseudomonas fluorescens
C:Species: Pseudomonas fluorescens
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C:Accession: S10527
R:Gilbert, H.J.; Hall, J.; Hazlewood, G.P.; Ferreira, L.M.A.
Mol. Microbiol. 4, 759-767, 1990
A:Title: The N-terminal region of an endoglucanase from Pseudomonas fluorescens subsp
A:Reference number: S10527; MUID:90355836; PMID:2117693
A:Accession: S10527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-511 <GIL>
A:Cross-references: EMBL:X52615; NID:945497; PIDN:CAA36844.1; PID:945498
C:Superfamily: bacterial cellulose-binding domain homology; glycosidase GMCW domain h
F:31128/Domain: bacterial cellulose-binding domain homology <BCB>
F:180-217/Domain: glycosidase GMCW domain homology <GMC>
F:32-127/Disulfide bonds: #status predicted

Query Match
Best Local Similarity 72.2%; Score 39; DB 2; Length 511;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QLPCPAEL 8
Db 485 EVPCPAEL 492

RESULT 10
A32244
60k cysteine-rich outer membrane protein 2 precursor, serotype L1 and L2 - Chlamydia
C:Species: Chlamydia trachomatis
C>Date: 12-Oct-1989 #sequence_revision 27-Jun-1994 #text_change 16-Jul-1999
C:Accession: A32244; A43584; A36043; A30472; J0419; S18981; S24277
R:Allen, J.E.; Stephens, R.S.
J. Bacteriol. 171, 285-291, 1989
A:Title: Identification by sequence analysis of two-site posttranslational processing
A:Reference number: A32244; MUID:89123030; PMID:2914847
A:Accession: A32244
A:Molecule type: DNA
A:Residues: 1-547 <ALL>
A:Cross-references: GB:M23001; NID:9144552; PIDN:AAA23152.1; PID:9144553

A:Experimental source: strain L2/434/Bu
 A:Note: parts of this sequence, including the amino ends of the precursor and mature proteins, are identical to the sequence of the 60-kDa outer membrane protein of Chlamydia trachomatis serovar L2.
 A:Reference number: S13120; MUID:91067486; PMID:2251143
 A:Accession: S13120
 A:Molecule type: DNA
 A:Residues: 1-32, 'F', 35-120, 'L', 122-131, 'A', 133-457, 'S', 459-547 <CON>
 A:Cross-references: EMBL:X55903; NID:940724; PIDN:CAA33936.1; PID:940725
 A:Experimental source: strain DK20, serotype E
 C:Genetics:
 A:Gene: omp2; omcB
 C:Function:
 A:Note: associated with differentiation of reticulate bodies into elementary bodies
 A:Superfamily: 60k cysteine-rich outer membrane protein
 C:Keywords: membrane protein; virulence
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-46/Domain: propeptide #status predicted <PRO>
 F:47-547/Product: 60k cysteine-rich outer membrane protein 2 #status predicted <MAT>

Query Match 72.2% Score 39; DB 1; Length 547;
 Best Local Similarity 70.0%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 148
 DB 139 QLPCEAEFVR 148

RESULT 11
 B43584
 60k cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
 C:Species: Chlamydia trachomatis
 C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 20-Aug-1999
 C:Accession: B43584; S13120; S18979; S24275
 Ride la Maza, L.M.; Felder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
 Infect. Immun. 59, 1196-1201, 1991
 A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein
 A:Reference number: A43584; MUID:91147205; PMID:1997423
 A:Accession: B43584
 A:Molecule type: DNA
 A:Residues: 1-547
 A:Cross-references: EMBL:X54389; NID:940763; PIDN:CAA38259.1; PID:940764
 A:Experimental source: Strain Bour, serotype E
 R:Coles, A.M.; Allan, I.; Pearce, J.H.

Nucleic Acids Res. 18, 6713, 1990
 A:Title: The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia trachomatis serovar L2.
 A:Reference number: S13120; MUID:91067486; PMID:2251143
 A:Accession: S13120
 A:Molecule type: DNA
 A:Residues: 1-32, 'F', 35-120, 'L', 122-131, 'A', 133-457, 'S', 459-547 <CON>
 A:Cross-references: EMBL:X55903; NID:940724; PIDN:CAA33936.1; PID:940725
 A:Experimental source: strain DK20, serotype E
 C:Genetics:
 A:Gene: omp2; omcB
 C:Function:
 A:Note: associated with differentiation of reticulate bodies into elementary bodies
 A:Superfamily: 60k cysteine-rich outer membrane protein
 C:Keywords: membrane protein; virulence
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-46/Domain: propeptide #status predicted <PRO>
 F:47-547/Product: 60k cysteine-rich outer membrane protein 2 #status predicted <MAT>

Query Match 72.2% Score 39; DB 2; Length 547;
 Best Local Similarity 70.0%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 148
 DB 139 QLPCEAEFVR 148

RESULT 12
 D71515
 60k cysteine-rich outer membrane protein 2 precursor serotypes B, C, and D - Chlamydia trachomatis
 C:Species: Chlamydia trachomatis
 C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 20-Aug-1999
 C:Accession: D71515; C43584; S11673
 R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
 A:Reference number: A71570; MUID:9900809; PMID:9784136
 A:Accession: D71515
 A:Molecule type: DNA
 A:Residues: 1-553 <ARN>
 A:Cross-references: GB:AEO01317; GB:AEO01273; NID:93328663; PIDN:AAC68042.1; PID:9332
 A:Experimental source: serotype D, strain UW-3/Cx
 Ride la Maza, L.M.; Felder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
 Infect. Immun. 59, 1196-1201, 1991
 A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein
 A:Reference number: A43584; MUID:91147205; PMID:1997423
 A:Accession: C43584
 A:Molecule type: DNA
 A:Residues: 7-238, 'V', 240-553
 A:Cross-references: GB:X54388; NID:940760; PIDN:CAA38257.1; PID:940761
 A:Experimental source: serotype C
 R:Watson, M.W.; Lambden, P.R.; Ward, M.E.; Clarke, I.N.
 FEMS Microbiol. Lett. 65, 293-297, 1989
 A:Title: Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein: sequence
 A:Reference number: S11673
 A:Accession: S11673
 A:Molecule type: DNA
 A:Residues: 7-553 <MAT>
 A:Cross-references: EMBL:X53510; NID:940681; PIDN:CAA37588.1; PID:940683
 A:Experimental source: serotype B
 C:Genetics:
 A:Gene: omp2; omcB
 C:Function:
 A:Note: associated with differentiation of reticulate bodies into elementary bodies
 A:Superfamily: 60k cysteine-rich outer membrane protein
 C:Keywords: membrane protein; virulence
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-46/Domain: propeptide #status predicted <PRO>
 F:47-553/Product: 60k cysteine-rich outer membrane protein 2 #status predicted <MAT>

Best Local Similarity 70.0%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QPCPAELLR 10
DB 145 QPCPAEFFVR 154

RESULT 13

E87644

sensory box histidine kinase/response regulator [imported] - *Caulobacter crescentus*C:Species: *Caulobacter crescentus*

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: E87644

R:Niernmen, W.C.; Feldbljum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolorn, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87644

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-642 <STO>

A:Cross-references: GB:AE005673; NID:q13424865; PIDN:AAK25153.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3191

Query Match

Best Local Similarity 70.4%; Score 38; DB 2; Length 642;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CPAPELLR 10
DB 340 CPAPELLR 346

RESULT 14

F83615

hypothetical protein PA0244 [imported] - *Pseudomonas aeruginosa* (strain PA01)C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: F83615

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: F83615

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-284 <STO>

A:Cross-references: GB:AE004462; GB:AE004091; NID:g9946077; PIDN:AA603633.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0244

C:Superfamily: shikimate dehydrogenase; shikimate dehydrogenase homology

Query Match

Best Local Similarity 68.5%; Score 37; DB 2; Length 284;

Best Local Similarity 77.8%; Pred. No. 36;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPCPAELLR 10
DB 211 MPVPAELLR 219

RESULT 15

G84331

hypothetical protein Vng1800h [imported] - *Halobacterium* sp. NRC-1C:Species: *Halobacterium* sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84331

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky

; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: G84331

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-63 <STO>

A:Cross-references: GB:AE004437; NID:q10581253; PIDN:AA620019.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG1800H

Query Match

Best Local Similarity 66.7%; Score 36; DB 2; Length 63;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAE 7
DB 18 LPCPAE 23

Search completed: July 9, 2003, 12:02:43
Job time : 20.5 secs

4
5

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:55:50 ; Search time 15.5 Seconds

(Without alignments)
26.759 Million cell updates/sec

Title: US-09-851-058-1

Sequence: 1 QLECPAELLR 10

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	54	100.0	471 1	TNAA_ECO57
2	54	100.0	471 1	TNAA_ECO57
3	54	100.0	471 1	TNAA_PASMU
4	54	100.0	472 1	TNAA_HAEIN
5	47	87.0	472 1	TNAA_VIRCH
6	47	75.9	445 1	ML64_MOUSE
7	41	75.9	446 1	ML64_MOUSE
8	41	75.9	446 1	ML64_MOUSE
9	39	72.2	438 1	CAC2_YEAST
10	39	72.2	511 1	RNH2_XYLFA
11	39	72.2	547 1	GUNB_PSEFL
12	39	72.2	547 1	GUNB_PSEFL
13	39	72.2	547 1	GUNB_PSEFL
14	39	72.2	547 1	GUNB_PSEFL
15	38	70.4	632 1	STC_CLOBE
16	37	68.5	547 1	STC_CLOBE
17	37	68.5	547 1	STC_CLOBE
18	36	66.7	370 1	YAF9_SCHPO
19	36	66.7	370 1	YAF9_SCHPO
20	36	66.7	370 1	YAF9_SCHPO
21	35	64.8	213 1	TBX1_HUMAN
22	35	64.8	257 1	GUN5_HUMAN
23	34	63.0	196 1	IDLC_HUMAN
24	34	63.0	222 1	HTGA_ECOLI
25	34	63.0	222 1	HTGA_ECOLI
26	34	63.0	222 1	HTGA_ECOLI
27	34	63.0	222 1	HTGA_ECOLI
28	34	63.0	222 1	HTGA_ECOLI
29	34	63.0	222 1	HTGA_ECOLI
30	33	61.1	156 1	CR18_ORYSA
31	33	61.1	274 1	AROE_PSEAE
32	33	61.1	295 1	ISPE_PASMU
33	33	61.1	361 1	IHA_TRIVU

34	33	61.1	365 1	P43_XENBO	P25066 xenopus bor
35	33	61.1	365 1	P43_XENBO	P25456 xenopus lae
36	33	61.1	370 1	TM61_HUMAN	09bzx5 homo sapien
37	33	61.1	397 1	LICH_MOUSE	09z0m5 mus musculu
38	33	61.1	448 1	EDAR_MOUSE	09unm0 homo sapien
39	33	61.1	448 1	EDAR_MOUSE	09unm0 homo sapien
40	33	61.1	511 1	U171_CAEEL	020069 caenorhabd1
41	33	61.1	532 1	PMGI_SYNY3	022757 synchocyst
42	33	61.1	560 1	YML2_ARATH	022757 arabidopsis
43	33	61.1	567 1	ENV_AVISN	P31796 avian splee
44	33	61.1	668 1	YFCK_ECOLI	P77182 escherichia
45	33	61.1	830 1	SREC_HUMAN	014162 homo sapien

ALIGNMENTS

RESULT 1
ID TNAA_ECO57 STANDARD: PRT: 471 AA.
AC Q8XB34;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Trypophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (Tnase).
GN TNAA OR 25203 OR EGS4645.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perina N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potomousis K.,
RA Anodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shigaawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
-1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> indole + pyruvate +
NH(3).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Tryptophan catabolism;
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY;
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL: AF005602; AAC58908.1; ALT_INT.
DR EMBL: AP002566; BAB38068.1; ALT_INT.
DR InterPro: IPR001597; Beta_elim_lyase.
DR Pfam: PF01212; Beta_elim_lyase; 1.
DR ProDom: PD005927; Beta_elim_lyase; 1.

DR PROSITE: P500853; BETA_ELM_LYASE. 1.
 KM Tryptophan catabolism; Lyase; PYRIDOXAL PHOSPHATE. Complete proteome.
 FT BINDING 270 270
 SQ SEQUENCE 471 AA; 52789 MW; 49DD2D41BD9D0034 CRC64;

Query Match 100.0%; Score 54; DB 1; Length 471;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLPCEALLR 10
 |||||
 Db 410 QLPCEALLR 419

RESULT 2
 ID TNAI_ECOLI STANDARD; PRT; 471 AA.
 AC P00913; P78123;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
 GN TNAI OR IND OR B3708.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCBI_TaxID:562;
 CX
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=82007678; PubMed=6268608;
 RA Dealey M.C., Yanofsky C.;
 RT "Nucleotide sequence of the structural gene for tryptophanase of
 Escherichia coli K-12.";
 RL J. Bacteriol. 147:787-796(1981).
 RN [12]
 RP SEQUENCE FROM N.A., AND IMPORTANCE OF CYS-298.
 RC STRAIN-B/17-A;
 RX MEDLINE=89333226; PubMed=2502187;
 RA Tsuchimizu M., Tsujimoto N., Oda T., Honda T., Yumoto N., Ito S.,
 RA Yamamoto M., Kim E.H., Hiragi Y.;
 RT "Role of cysteine residues in tryptophanase for monovalent cation-
 induced activation.";
 RL Biochimie 71:711-720(1989).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=93315143; PubMed=7668882;
 RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
 RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
 genome: organizational symmetry around the origin of replication.";
 RL Genome 16:551-561(1993).
 RN [4]
 RP SEQUENCE OF TRYPTIC PEPTIDES.
 RC STRAIN-K12;
 RX MEDLINE=7214434; PubMed=4551944;
 RA Kagamiyama H., Matsubara H., Snell E.E.;
 RT "The chemical structure of tryptophanase from Escherichia coli. 3.
 Isolation and amino acid sequence of the tryptic peptides.";
 RL J. Biol. Chem. 247:1576-1586(1972).
 RN [5]
 RP SEQUENCE OF 1-21 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=86033634; PubMed=3902796;
 RA Stewart V., Yanofsky C.;
 RT "Evidence for transcription antitermination control of tryptophanase
 operon expression in Escherichia coli K-12.";
 RL J. Bacteriol. 164:731-740(1985).
 RN [6]
 RP SEQUENCE OF 463-471 FROM N.A.
 RX MEDLINE=91216998; PubMed=2022620;
 RA Sarsoro J.P., Wooley P.J., Gollnick P.D., Yanofsky C., Pittard A.J.;
 RT "A new family of integral membrane proteins involved in transport of

RT aromatic amino acids in Escherichia coli.";
 RL J. Bacteriol. 173:3231-3234(1991).
 RN [7]
 RP SEQUENCE OF 1-12.
 RC STRAIN-K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 in the genome of Escherichia coli K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 RN [8]
 RP MUTAGENESIS OF CYS-294 AND CYS-298.
 RX MEDLINE=89278130; PubMed=2659590;
 RA Phillips R.S., Gollnick P.D.;
 RT "Evidence that cysteine 298 is in the active site of tryptophan
 indole-lyase.";
 RL J. Biol. Chem. 264:10627-10632(1989).
 CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O = indole + pyruvate +
 NH(3).
 CC -1- CORFACTOR: Pyridoxal phosphate.
 CC -1- PATHWAY: Tryptophan catabolism.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as their content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: K00032; AA24676.1; -;
 DR EMBL: X15974; CA434096.1; -;
 DR EMBL: M15900; AA24679.1; -;
 DR EMBL: M59914; -; NOT_ANNOTATED_CDS.
 DR EMBL: L10328; AA62059.1; ALT_INIT.
 DR EMBL: AE000448; AAC76731.1; ALT_INIT.
 DR PIR: A01136; WZEC.
 DR HSSP: P28796; 1AX4.
 DR EC02DBASE: G046.5; 6TH EDITION.
 DR Ecogene: E011005; tnaA.
 DR InterPro: IPR001597; Beta_elm_lyase.
 DR Pfam: PF01212; Beta_elm_lyase.1.
 DR ProDom: PD005927; Beta_elm_lyase.1.
 DR PROSITE: P500853; BETA_ELM_LYASE.1.
 KM Tryptophan catabolism; Lyase; PYRIDOXAL PHOSPHATE. Complete proteome.
 FT BINDING 270 270
 FT MUTAGEN 294 294 C->S: IDENTICAL TO WILD-TYPE.
 FT MUTAGEN 298 298 C->S: ALTERED ACTIVITY.
 FT CONFLICT 137 140 DTG -> TDOG (IN REF. 1).
 FT CONFLICT 379 380 QA -> TG (IN REF. 1 AND 2).
 SQ SEQUENCE 471 AA; 52773 MW; 5AFC1F41BD9D0034 CRC64;

Query Match 100.0%; Score 54; DB 1; Length 471;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLPCEALLR 10
 |||||
 Db 410 QLPCEALLR 419

RESULT 3
 ID TNAI_PASMU STANDARD; PRT; 471 AA.
 AC Q9CL27;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
 GN TNAI OR PM1420.
 OS Pasteurella multocida.

```

OC Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
RT "Complete genome sequence of Pasteurella multocida PM70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> Indole + pyruvate +
CC NH(3).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Tryptophan catabolism.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006178; AAK03504.1; -.
DR HSSP: P28796; IAX4.
DR InterPro: IPR001597; Beta_elim_lyase.
DR Pfam: PF01212; Beta_elim_lyase; 1.
DR ProDom: PD005927; Beta_elim_lyase; 1.
DR PROSITE: PS00853; BETA_ELIM_LYASE; 1.
KW Tryptophan catabolism; Lyase; Pyridoxal phosphate; Complete proteome.
FT BINDING 270 270 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 471 AA; 52893 MW; B6DE142FF1A7BF8E CRC64;

Query Match 100.0%; Score 54; DB 1; Length 471;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCEALLR 10
DB 410 QLPCEALLR 419

RESULT 4
TAXA_VIBCH STANDARD; PRT; 472 AA.
ID TAXA_VIBCH
AC 007674;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
OS TNase.
GN Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Esagan / Serotype B;
RX MEDLINE=98083063; PubMed=9422600;
RA Martin K., Morlin G., Smith A., Nordyke A., Eisenstark A., Golomb M.;
RT "The tryptophanase gene cluster of Haemophilus influenzae type b:
RT evidence for horizontal gene transfer.";
RL J. Bacteriol. 180:107-116(1998).
CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> Indole + pyruvate +
CC NH(3).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Tryptophan catabolism.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- MISCELLANEOUS: Some pathogenic strains of H. influenzae are active
CC in tryptophan catabolism and contains tna genes which appear to
CC have been inserted as a mobile unit.

```

```

CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF003252; AAB96579.1; -.
DR HSSP: P28796; IAX4.
DR InterPro: IPR001597; Beta_elim_lyase.
DR Pfam: PF01212; Beta_elim_lyase; 1.
DR ProDom: PD005927; Beta_elim_lyase; 1.
DR PROSITE: PS00853; BETA_ELIM_LYASE; 1.
KW Tryptophan catabolism; Lyase; Pyridoxal phosphate.
FT BINDING 270 270 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 472 AA; 53088 MW; 8DE063F4D89F04E2 CRC64;

Query Match 100.0%; Score 54; DB 1; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCEALLR 10
DB 410 QLPCEALLR 419

RESULT 5
TAXA_VIBCH STANDARD; PRT; 472 AA.
ID TAXA_VIBCH
AC 09KN05; 086039;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
GN TNase OR VCA0161.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eise J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Hart D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
RA McDonald L., Uitterlind T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
RN [2]
RP SEQUENCE OF 51-241 FROM N.A.
RC STRAIN=Bengal;
RA Tang T.H., Ravichandran M., Johari M.R., Zainuddin Z.F.;
RT "Vibrio cholerae putative tryptophanase gene partial cds.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> Indole + pyruvate +
CC NH(3).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Tryptophan catabolism.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE004357; AAF6074.1; -

DR EMBL: AF081274; AAC3284.1; -

DR HSSP: P28796; 1AX4.

DR TIGR: VCA0161; -

DR InterPro: IPR001597; Beta_elim_lyase.

DR Pfam: PF01212; Beta_elim_lyase; 1.

DR PRODOM: PD005927; Beta_elim_lyase; 1.

DR PROSITE: PS00853; BETA_ELIM_LYASE; 1.

KW Tryptophan catabolism; Lyase; pyridoxal phosphate; Complete proteome.

FT BINDING 270 270 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

SO SEQUENCE 472 AA; 52920 MW; 632384F54C06F7B9 CRC64;

Query Match

Best Local Similarity 87.0%; Score 47; DB 1; Length 472;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 QLECPAEELR 10

DB 410 QHCPAEELR 419

RESULT 6

ML64_HUMAN STANDARD; PRT; 445 AA.

AC Q14849; Q96HM9; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE MLN 64 protein (CABL protein) (Stard3).

GN MLN64 OR CAB1 OR STARD3.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Breast carcinoma;

RX MEDLINE-96039245; PubMed-7490069;

RA Tomasetto C., Regnier C., Moog-Lutz C., Mattei M.-G., Chenard M.-P., Lidenau R., Bassot P., Rio M.-C.;

RT "Identification of four novel human genes amplified and overexpressed in breast carcinoma and localized to the q11-q21.3 region of chromosome 17.";

RT Chromosome 17.

RL Genomics 28:367-376(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-Oesophageal carcinoma;

RX MEDLINE-97413641; PubMed-9270027;

RA Akiyama N., Sasaki H., Ishizuka T., Kishi T., Sakamoto H., Onda M., Hirai H., Yazaki Y., Sugimura T., Terada M.;

RT "Identification of a candidate gene, CAB1, for cholesterol transport to mitochondria from the c-erbB-2 amplicon by a modified cDNA selection method.";

RT Cancer Res. 57:3548-3553(1997).

RL [3]

RP SEQUENCE FROM N.A.

RC TISSUE-Lung, Skin, and Spleen;

RA Strausberg R.;

RP Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 216-445.

RX MEDLINE-20264523; PubMed-10802740;

RA Tanihata Y., Hurley J.H.;

RT "Structure and lipid transport mechanism of a STAR-related domain.";

RL Nat. Struct. Biol. 7:408-414(2000).

CC -1- FUNCTION: Binds and transports cholesterol. Promotes steroidogenesis in placenta and brain.

CC -1- SIMILARITY: CONTAINS 1 START DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X80199; CA56489.1; -

DR EMBL: D38253; BAA22525.1; -

DR EMBL: BC008356; AAH08356.1; -

DR EMBL: BC008747; AAH08747.1; -

DR EMBL: BC025679; AAH25679.1; -

DR PDB: 1EM2; 02-MAY-00.

DR MIM: 607048; -

DR InterPro: IPR002913; START.

DR InterPro: IPR00799; STAR.

DR Pfam: PF01852; START; 1.

DR PRINTS: PR00978; STARPROTEIN.

DR SMART: SM00234; START; 1.

DR PROSITE: PS00848; START; 1.

KW Lipid-binding; Lipid transport; Transport; Steroidogenesis;

KW 3D-structure.

FT DOMAIN 230 443 START.

FT CONFLICT 117 117 G -> R (IN REF. 3: AAH08356/AAH25679).

FT CONFLICT 216 216 G -> A (IN REF. 3: AAH25679).

SO SEQUENCE 445 AA; 50474 MW; 62BBD5C3EAD0DEF CRC64;

Query Match

Best Local Similarity 75.9%; Score 41; DB 1; Length 445;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAEL 9

DB 289 LPCPAELV 296

RESULT 7

ML64_MOUSE STANDARD; PRT; 446 AA.

AC O61542; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE MLN 64 protein (ES 64 protein) (Stard3).

GN MLN64 OR ES64 OR STARD3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

OX [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Breast;

RX MEDLINE-96039245; PubMed-7490069;

RA Tomasetto C., Regnier C.H., Moog-Lutz C., Mattei M.-G., Chenard M.-P., Lidenau R., Bassot P., Rio M.-C.;

RT "Identification of four novel human genes amplified and overexpressed in breast carcinoma and localized to the q11-q21.3 region of chromosome 17.";

RT Chromosome 17.

RL Genomics 28:367-376(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-Breast;

RA Strausberg R.;

RP Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RX MEDLINE-20264523; PubMed-10802740;

RA Tanihata Y., Hurley J.H.;

RT "Structure and lipid transport mechanism of a STAR-related domain.";

RL Nat. Struct. Biol. 7:408-414(2000).

CC -1- FUNCTION: Binds and transports cholesterol. Promotes steroidogenesis in placenta and brain (By similarity).

CC -1- SIMILARITY: CONTAINS 1 START DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch.
CC -----
DR EMBL: X82457; CAA57834.1; -
DR EMBL: BC003313; AAH03313.1; -
DR HSSP: Q14849; IEM2. -
DR MGD: MGI:1929618; Minc4. -
DR InterPro: IPR002913; START.
DR InterPro: IPR000799; STAR.
DR Pfam: PF01852; START. 1.
DR PRINTS: PR00878; STARPROTEIN.
DR SMART: SM00234; START. 1.
DR PROSITE: PS50848; START. 1.
DR Lipid-binding: Lipid transport; Transport; Steroidogenesis.
DR DOMAIN 231 444 START.
DR SEQUENCE 446 AA; 50469 MW; DBF4359604F3E1E2 CRC64;
SO
Query Match 75.98; Score 41; DB 1; Length 446;
Best Local Similarity 87.58; Pred. No. 4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 2 LKCPAEIL 9
Db 290 LKCPAEIL 297
RESULT 8
CAC2_YEAST
ID CAC2_YEAST STANDARD: PRT: 468 AA.
AC 004199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Chromatin assembly factor 1 p60 subunit (CAF-1 60 kDa subunit).
GN CAC2 OR WIL102W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN-S288C; AB972;
RA Devlin K., Churcher C., Bartell B.G., Rajandream M.A.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
RT [2]
RP SEQUENCE OF 58-83 AND 459-468, AND CHARACTERIZATION.
RX MEDLINE-97182548; PubMed-9030687.
RA Kaufman P.D., Kobayashi R., Stillman B.;
RT "Ultraviolet radiation sensitivity and reduction of telomeric
silencing in Saccharomyces cerevisiae cells lacking chromatin
assembly factor-1."
RL Genes Dev. 11:345-357(1997).
CC -1- FUNCTION: COMPLEX THAT ASSEMBLES HISTONE OCTAMERS ONTO REPLICATING
CC DNA IN VITRO. CAF-1 PERFORMS THE FIRST STEP OF THE NUCLEOSOME
CC ASSEMBLY PROCESS, BRINGING NEWLY SYNTHESIZED HISTONES H3 AND H4 TO
CC REPLICATING DNA. HISTONES H2A/H2B CAN BIND TO THIS CHROMATIN
CC PRECURSOR SUBSEQUENT TO DNA REPLICATION TO COMPLETE THE HISTONE
CC OCTAMER. P150 AND P60 FORM COMPLEXES WITH NEWLY SYNTHESIZED
CC HISTONES H3 AND ACETYLATED H4 IN CELL EXTRACTS (BY SIMILARITY).
CC -1- SUBUNIT: COMPOSED OF THREE SUBUNTS, P50, P60 AND P90.
CC -1- SUBCELLULAR LOCATION: Nucleus.
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE HIR1 FAMILY OF WD-REPEAT PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X80835; CAA56795.1; -
DR SGD: S0004570; CAC2.

DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 5.
DR PRINTS: PR00320; GPROTEINRPT.
DR SMART: SM00320; WD40; 5.
DR PROSITE: PS00678; WD_REPEATS_1; FALSE_NEG.
DR PROSITE: PS50082; WD_REPEATS_2; 2.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT REPEAT 11 52 WD 1.
FT REPEAT 69 108 WD 2.
FT REPEAT 143 182 WD 3.
FT REPEAT 185 224 WD 4.
FT REPEAT 371 413 WD 5.
SO SEQUENCE 468 AA; 51252 MW; F24853C4AAA4D97C2 CRC64;
Query Match 75.98; Score 41; DB 1; Length 468;
Best Local Similarity 60.08; Pred. No. 4.2;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Oy 1 OLKCPAEILR 10
Db 237 ELKCPGVLR 246
RESULT 9
RNH2_XYLFA
ID RNH2_XYLFA STANDARD: PRT: 234 AA.
AC Q9PEI7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease H1I (EC 3.1.26.4) (RNase H1I).
GN RNHB OR XR1041.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN-9a5c;
RX MEDLINE-20365717; PubMed-10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franco S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.V., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Trufi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidans J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -1- FUNCTION: THIS ENZYME IS AN ENDONUCLEASE THAT DEGRADES THE RNA OF
CC RNA-DNA HYBRIDS SPECIFICALLY (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.

CC -1- COFACTOR: MANGANESE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE RNASE HII FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AEO03941; AAF83851.1;
 DR InterPro: IPR001352; RNase_HII/HIII.
 DR Pfam: PF01351; RNase_HII.1.
 KM Hydrolase; Nuclease; Endonuclease; Manganese; Complete proteome.
 FT ACT_SITE 22 22 BY SIMILARITY.
 FT ACT_SITE 115 115 BY SIMILARITY.
 FT ACT_SITE 134 134 BY SIMILARITY.
 SQ SEQUENCE 234 AA; 25633 MW; DFB4EADBE145F22 CRC64;

Query Match 72.2%; Score 39; DB 1; Length 234;
 Best Local Similarity 87.5%; Pred. No. 5;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPCPAEL 9
 DB 123 LPCPAEL 130

RESULT 10
 GUNB_PSEFL STANDARD: PRT; 511 AA.
 AC P18126;
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Endoglucanase B precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
 DE (Cellulase) (ECB).
 GN CELB.
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=294;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-48.
 RC STRAIN=SP. Cellulosa;
 RX MEDLINE=90355836; PubMed=2117693;
 RA Gilbert H.J., Hall J., Hazlewood G.P., Ferreira L.M.A.;
 RT The N-terminal region of an endoglucanase from Pseudomonas
 RT fluorescens subspecies cellulosa constitutes a cellulose-binding
 RT domain that is distinct from the catalytic centre.*;
 RL Mol. Microbiol. 4:759-767(1990).
 CC -1- FUNCTION: THIS ENZYME CATALYZES THE ENDOHYDROLYSIS OF 1,4-BETA-
 CC GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-D-
 CC GLUCANS. ECB IS MOST ACTIVE AGAINST BARLEY BETA-GLUCAN, BUT SHOWED
 CC SIGNIFICANT ACTIVITY AGAINST AMORPHOUS AND CRYSTALLINE CELLULOSE.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: CONTAINS 1 BACTERIAL-TYPE CELLULOSE-BINDING DOMAIN
 CC (CBD).
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY K (FAMILY 45 OF GLYCOSYL
 CC HYDROLASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X52615; CAA36844.1;
 DR PIR: S10527; S10527.
 DR HSSP: P43316; 2ENG.
 DR InterPro: IPR001919; Bac_celose-blnd.
 DR InterPro: IPR002883; CBD_5.
 DR InterPro: IPR000334; GH_45.
 DR Pfam: PF00553; CBM_2; 1.
 DR Pfam: PF02013; CBM_10; 1.
 DR Pfam: PF02015; Glyco_hydro_45; 1.
 DR PROSITE: PS00561; CBD_BACTERIAL; 1.
 DR PROSITE: PS01140; GLYCOSYL_HYDROL_F45; 1.
 DR Cellulose degradation; Hydrolase; Glycosidase; Signal; Periplasmic.
 FT SIGNAL 1 29
 FT CHAIN 30 511
 FT DOMAIN 30 131
 FT DOMAIN 132 173
 FT DOMAIN 223 259
 FT DISULFID 32 127
 FT ACT_SITE 276 276 BY SIMILARITY.
 FT ACT_SITE 393 393 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 393 393 PROTON DONOR (BY SIMILARITY).
 SQ SEQUENCE 511 AA; 52078 MW; 3C3119D998291D8E CRC64;

Query Match 72.2%; Score 39; DB 1; Length 511;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 OLPCPAEL 8
 DB 485 EVPCPAEL 492

RESULT 11
 OM6C_CHLTR STANDARD: PRT; 547 AA.
 AC P26758;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 60 kDa outer membrane protein, serovar C precursor (Cysteine-rich
 DE outer membrane protein) (60-kDa CRP).
 GN OMP2 OR OMP2B.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C/TW-3;
 RX MEDLINE=91147205; PubMed=1997423;
 RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,
 RA Peterson E.M.;
 RT *Sequence diversity of the 60-kilodalton protein and of a putative
 RT 15-kilodalton protein between the trachoma and lymphogranuloma
 RT venereum biovars of Chlamydia trachomatis.*;
 RL Infect. Immun. 59:1196-1201(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zhang Y.X., Caldwell H.D.;
 RA Subling Y.H. (1992) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULAR BODIES
 CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
 CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
 CC VIRULENCE FACTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X54388; CAA38257.1;

DR EMBL: M85197; AAA23159.1; -.
DR PIR: C43584; C43584.
DR InterPro: IPR003506; Chlam_OMP6.
DR Pfam: PF03504; Chlam_OMP6.1.
DR PRINTS: PR01336; CHLAMIDIOM6.
KW Outer membrane; Transmembrane; Signal; Virulence.
FT SIGNAL 1 22
FT PROPEP 23 40
FT CHAIN 41 547
SQ SEQUENCE 547 AA; 58680 MM; 817BA5DCEFE6504 CRC64;
Query Match 72.2% Score 39; DB 1; Length 547;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 QLPCEAEFLR 10
Db 139 QLPCEAEFVR 148
RESULT 12
OM6D_CHLTR STANDARD; PRT; 547 AA.
ID OM6D_CHLTR
AC P18151;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane protein) (CRP) (60 kDa cysteine-rich OMP).
GN OMCB OR OMP2 OR OMP2B OR CT443.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B/Jal120/OT;
RX MEDLINE=90128208; PubMed=2612891;
RA Watson M.W., Lambden P.R., Ward M.E., Clarke I.N.;
RT "Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein: sequence homology between trachoma and LGV biovars.";
RL FEBS Microbiol. Lett. 53:293-297(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-B/Tw-5/OT;
RX MEDLINE=91141306; PubMed=2287277;
RA Allen J.E., Cerrone M.C., Beatty P.R., Stephens R.S.;
RT "Cysteine-rich outer membrane proteins of Chlamydia trachomatis display compensatory sequence changes between biovariants.";
RL Mol. Microbiol. 4:1543-1550(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-D/W-3/CX;
RX MEDLINE=99008009; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL: X53510; CAA37588.1; -.
DR EMBL: AE001317; AAC68042.1; ALT_INIT.
DR PIR: S11673; S11673.
DR PRC1-2DPAGE; P18151; -.
DR InterPro: IPR003506; Chlam_OMP6.
DR Pfam: PF03504; Chlam_OMP6.1.
DR PRINTS: PR01336; CHLAMIDIOM6.
KW Outer membrane; Transmembrane; Signal; Virulence; Complete proteome.
FT SIGNAL 1 22
FT PROPEP 23 40
FT CHAIN 41 547
FT VARIANT 233 233
SQ SEQUENCE 547 AA; 58694 MM; 42719B4BCECC6A CRC64;
Query Match 72.2% Score 39; DB 1; Length 547;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 QLPCEAEFLR 10
Db 139 QLPCEAEFVR 148
RESULT 13
OM6E_CHLTR STANDARD; PRT; 547 AA.
ID OM6E_CHLTR
AC P23603;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein, serovar E precursor (Cysteine-rich outer membrane protein) (60-kDa CRP).
GN OMP2 OR OMP2B.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BOR / Serovar E;
RX MEDLINE=91067486; PubMed=2251143;
RA Coles A.M., Allan I., Pearce J.H.;
RT "The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia trachomatis serovar E.";
RL Nucleic Acids Res. 18:6713-6713(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BOR / Serovar E;
RX MEDLINE=91147205; PubMed=1997423;
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A., Peterson E.M.;
RT "Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein between the trachoma and lymphogranuloma venereum biovars of Chlamydia trachomatis.";
RL Infect. Immun. 59:1196-1201(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC Zhang Y.X., Caldwell H.D.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: X55903; CAA39396.1; -
DR EMBL: X54389; CAA38259.1; -
DR EMBL: M85196; AAA23154.1; -
DR PIR: S13120; S13120; -
DR InterPro: IPR003506; Chlam_OMP6.
DR Pfam: PF03504; Chlam_OMP; 1.
DR PRINTS: PR01336; CHLAMIDIAOM6.
KW Outer membrane; Transmembrane; Signal; Virulence.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 1 40
FT CHAIN 41 547
FT VARIANT 33 34 60 KDA OUTER MEMBRANE PROTEIN, SEROVAR E.
FT VARIANT 121 121 SL -> FT (IN STRAIN DK20).
FT VARIANT 132 132 I -> L (IN STRAIN DK20).
FT VARIANT 458 458 V -> A (IN STRAIN DK20).
FT VARIANT 458 458 N -> S (IN STRAIN DK20).
SQ SEQUENCE 547 AA; 58708 MW; 05206084F4E20AB CRC64;

Query Match 72.2%; Score 39; DB 1; Length 547;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAFELR 10
DB 139 QLPCEAFVR 148

RESULT 14
OM6L.CHLTR STANDARD; PRT: 547 AA.
AC P21354; P18586;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60 kDa outer membrane protein, serovars L1/L2/L3 precursor (Cysteine-
rich outer membrane protein) (60-kDa CRP).
GN omp2 OR omp2B.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=L2/434/Bu;
RA MEDLINE=89123030; PubMed=2914847;
RA Allen J.E., Stephens R.S.;
RT Identification by sequence analysis of two-site posttranslational
RT processing of the cysteine-rich outer membrane protein 2 of Chlamydia
RT trachomatis serovar L2."
RL J. Bacteriol. 171:285-291(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=L3/404;
RA MEDLINE=91147205; PubMed=1997423;
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,
RA Peterson E.M.;
RT *Sequence diversity of the 60-kilodalton protein and of a putative
RT 15-kilodalton protein between the trachoma and lymphogranuloma
RT venereum blavars of Chlamydia trachomatis."
RL Infect. Immun. 59:1196-1201(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=L1/440/LN;
RA MEDLINE=89138006; PubMed=3066701;
RA Clarke I.N., Ward M.E., Lambden P.R.;
RT "Molecular cloning and sequence analysis of a developmentally
RT regulated cysteine-rich outer membrane protein from Chlamydia
RT trachomatis."
RL Gene 71:307-314(1988).
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
CC VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23001; AAA23151.1; -
DR EMBL: X54390; CAA38261.1; -
DR EMBL: M5148; AAA23119.1; ALT_INIT.
DR PIR: A32244; A32244.
DR PIR: A32244; A32244.
DR PIR: S18981; S18981.
DR PIR: A43584; A43584.
DR PIR: J00419; J00419.
DR Slenz-2DPAGE; P21354; -
DR InterPro: IPR003506; Chlam_OMP6.
DR Pfam: PF03504; Chlam_OMP; 1.
DR PRINTS: PR01336; CHLAMIDIAOM6.
KW Outer membrane; Transmembrane; Signal; Virulence.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 1 40
FT CHAIN 41 547
FT CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN, SEROVAR E.
SQ SEQUENCE 547 AA; 58782 MW; 78CEB41CCE98472D CRC64;

Query Match 72.2%; Score 39; DB 1; Length 547;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAFELR 10
DB 139 QLPCEAFVR 148

RESULT 15
STC_CLOBE STANDARD; PRT: 632 AA.
AC P26047; Q9L9X1;
DT 01-MAY-1992 (Rel. 22, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Signal-transduction and transcriptional-control protein.
GN STC.
OS Clostridium beijerinckii (Clostridium MP).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1520;
RN RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B593;
RA Rifeat M.M., Chen J.S.;
RT "The predicted protein product of a putative regulatory gene from
RT Clostridium beijerinckii NRRL B593 is homologous to the NtrC subfamily
RT of transcriptional regulators."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: THE N-TERMINAL REGION CONTAINS A SIGMA-54 FACTOR
CC INTERACTION ATP-BINDING DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF157307; AAD54950.1; -
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR002197; HTH_Fts.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR002078; S1954_Interact.

```


DR Pfam; PF00158; Sigma54_activat; 1.
 DR Pfam; PF00989; PAS; 1.
 DR Pfam; PF02954; HTH_8; 1.
 DR SMART; SM00382; AAA; 1.
 DR SMART; SM00091; PAS; 1.
 DR TIGRFAMS; TIGR00229; sensory_box; 1.
 DR TIGRFAMS; TIGR01199; HTH_fls; 1.
 DR PROSITE; PS50112; PAS; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
 DR PROSITE; PS00676; SIGMA54_INTERACT_2; 1.
 DR PROSITE; PS00688; SIGMA54_INTERACT_3; 1.
 DR PROSITE; PS50045; SIGMA54_INTERACT_4; 1.
 KW Transcription regulation; Sensory transduction; DNA-binding;
 ATP-binding.
 KM DOMAIN 197 270 PAS.
 FT DOMAIN 324 554 SIGMA-54 FACTOR INTERACTION (POTENTIAL).
 FT NP_BIND 352 359 ATP (POTENTIAL).
 FT NP_BIND 416 425 ATP (POTENTIAL).
 FT DNA_BIND 606 625 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 632 AA; 71537 MW; C52A3D906B0B2CA CRC64;

Query Match 70.4%; Score 38; DB 1; Length 632;
 Best Local Similarity 55.6%; Pred. No. 19;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPCPAELLR 10
 :|||::||
 DB 27 IPCPSKIIR 35

Search completed: July 9, 2003, 12:00:43
 Job time : 17.5 secs

This Page Blank (exptb)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:57:21 ; Search time 35.5 seconds
(without alignments)
36.041 Million cell updates/sec

Title: US-09-851-058-1
54
Sequence: 1 OLPCPAELLR 10

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	74.1	57	4 Q96IH8	Q96IH8 homo sapien
2	40	74.1	336	10 Q9MAA9	Q9MAA9 arabidopsis
3	40	74.1	5149	16 Q9I179	Q9I179 pseudomonas
4	39	72.2	511	2 Q93FV8	Q93FV8 chlamydia t
5	39	72.2	547	2 Q93317	Q93317 chlamydia t
6	38	70.4	147	10 Q9LEB7	Q9LEB7 petroselinu
7	38	70.4	265	16 Q9P7J3	Q9P7J3 chlamydia m
8	38	70.4	642	16 Q9A3L2	Q9A3L2 caulobacter
9	38	70.4	1339	13 Q9W612	Q9W612 pimephales
10	38	70.4	1353	4 Q9Y4B5	Q9Y4B5 homo sapien
11	38	70.4	1362	13 Q90YH8	Q90YH8 brachydanio
12	37	68.5	284	16 Q916P4	Q916P4 pseudomonas
13	37	68.5	348	8 Q94W74	Q94W74 scarlatia
14	37	68.5	405	8 Q9NFC7	Q9NFC7 drosophila
15	37	68.5	989	4 Q94909	Q94909 homo sapien
16	36	66.7	63	17 Q9HP51	Q9HP51 halobacteri

17	36	66.7	286	12 Q8VZK4	Q8VZK4 camelopox v1
18	36	66.7	335	5 Q18103	Q18103 caenorhabdi
19	36	66.7	415	16 Q9HZU0	Q9HZU0 pseudomonas
20	36	66.7	504	11 Q9JHE5	Q9JHE5 rattus norv
21	36	66.7	504	11 Q9J188	Q9J188 rattus norv
22	36	66.7	563	12 Q8QM02	Q8QM02 compox viru
23	36	66.7	564	12 Q8Q016	Q8Q016 camelopox v1
24	36	66.7	604	2 Q9X619	Q9X619 streptomyce
25	36	66.7	693	10 Q8S705	Q8S705 oryza sativ.
26	36	66.7	890	5 Q9VGJ7	Q9VGJ7 drosophila
27	35.5	65.7	453	11 Q8VHS5	Q8VHS5 mus musculu
28	35	64.8	118	2 Q49829	Q49829 mycobacteri
29	35	64.8	146	10 Q9LZP8	Q9LZP8 arabidopsis
30	35	64.8	150	4 Q95514	Q95514 homo sapien
31	35	64.8	207	11 Q9D7P7	Q9D7P7 mus musculu
32	35	64.8	233	10 Q94GK5	Q94GK5 oryza sativ
33	35	64.8	281	16 Q9F492	Q9F492 enterobacte
34	35	64.8	291	2 Q9F492	Q9F492 rhizobium m
35	35	64.8	305	3 Q93782	Q93782 humicola gr
36	35	64.8	442	4 Q9H0M8	Q9H0M8 homo sapien
37	35	64.8	444	17 Q8TVU0	Q8TVU0 methanopyru
38	35	64.8	493	13 Q90M05	Q90M05 platichlys
39	35	64.8	493	13 Q90M00	Q90M00 pleurotome
40	35	64.8	550	3 Q9P5P9	Q9P5P9 neurospora
41	35	64.8	579	3 Q8X135	Q8X135 gibberella
42	35	64.8	579	3 Q8X134	Q8X134 gibberella
43	35	64.8	606	3 Q8C114	Q8C114 fusarium sp
44	35	64.8	622	4 Q9H9T2	Q9H9T2 homo sapien
45	35	64.8	646	6 Q95J15	Q95J15 macaca fasc

ALIGNMENTS

RESULT 1					
Q96IH8					
ID Q96IH8	PRELIMINARY:	PRT:	57 AA.		
AC Q96IH8:					
DT 01-DEC-2001 (TREMBLrel. 19, Created)					
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)					
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)					
DE Similar to erythroid differentiation and denucleation factor					
DE 1.					
OS Homo sapiens (Human).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
OX NCBI_TaxID=9606;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=SKIN;					
RA Strausberg R.;					
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.					
DR EMBL; BC007528; AA07528.1; -					
SQ SEQUENCE 57 AA; 6004 MW; 2132EA21A4311176 CRC64;					
Query Match	74.1%	Score 40;	DB 4;	Length 57;	
Best Local Similarity	66.7%;	Pred. NO. 2.2;			
Matches	6;	Conservative	2;	Mismatches	1;
				Indels	0;
				Gaps	0;
QY 2 LPCPAELLR 10					
DB 26 LPCPEDILR 34					
RESULT 2					
Q9MAA9					
ID Q9MAA9	PRELIMINARY:	PRT:	336 AA.		
AC Q9MAA9:					
DT 01-OCT-2000 (TREMBLrel. 15, Created)					
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)					
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)					
DE T12H1.6 protein (Hypothetical 37.7 kDa protein).					
GN T12H1.6 OR A73G05100.					

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 RX NCBL_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 Rong H., Koo H., Fujii C.Y., Utechtack T.R., Barnstead M.E.,
 Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC T12H1 genomic sequence.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
 Lee J.M., Onodera C.S., Quach H.L., Tang C., Tortum M., Wu H.C.,
 Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 Hayashizaki Y., Ishida J., Jones T., Kameya A., Karlin-Neumann G.,
 Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,
 Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
 Theologis A.;
 RT "Full length cDNA of gene At3g05100 (GI:15229693).";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 Tang C., Tortum M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 Kameya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
 Davis R.W., Ecker J.R., Theologis A.;
 RT "Arabidopsis Open Reading Frame (ORF) Clones.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC009177; AAF27016.1; -;
 DR EMBL: AY070370; AAL49688.1; -;
 DR EMBL: AY091328; AAL14267.1; -;
 DR InterPro: IPR000051; SAM_Bind.
 KW Hydrophobic protein.
 SQ SEQUENCE 336 AA: 37719 MW: 6FA910F7B4B2FA85 CRC64;

Query Match 74.1%; Score 40; DB 10; Length 336;
 Best Local Similarity 87.5%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPPCPAELL 9
 Db 137 LPPCPAELL 144

RESULT 3
 091179 PRELIMINARY: PRT; 5149 AA.
 AC 091179;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Probable non-ribosomal peptide synthetase.
 GN PA2402.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 RX NCBL_TaxID=287;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Ervin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Gardner R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004667; AAC05790.1; -;
 DR HSPB: P14687; 1AMU.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; Condensatn.
 DR InterPro: IPR000577; FGy kin.
 DR InterPro: IPR003880; Pantine_attach.
 DR Pfam: PF00501; AMP-binding; 4.
 DR Pfam: PF00668; Condensation; 6.
 DR Pfam: PF00550; PP-binding; 4.
 DR PROSITE: PS00075; ACP_DOMAIN; 4.
 DR PROSITE: PS00445; AMP_BINDING; 4.
 DR PROSITE: PS00445; FGy_KINASES; 2; UNKNOWN.1.
 DR PROSITE: PS00012; PHOSPHOTRANSFERIN; UNKNOWN.3.
 KW Phosphotransferin; Complete proteome.
 SQ SEQUENCE 5149 AA: 569179 MW: 20575B7618921CA0 CRC64;

Query Match 74.1%; Score 40; DB 16; Length 5149;
 Best Local Similarity 70.0%; Pred. No. 14e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 10
 Db 2098 QLPCEAEFVR 2107

RESULT 4
 093FV8 PRELIMINARY: PRT; 511 AA.
 AC 093FV8;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE 60 kDa cysteine-rich OMP (Fragment).
 GN OMCB.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RX NCBL_TaxID=813;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAR-13;
 RX MEDLINE=21450826; PubMed=11567000;
 RA Millman K.L., Tavare S., Dean D.;
 RT "Recombination in the ompA gene but not the omcB gene of Chlamydia
 RT contributes to serovar-specific differences in tissue tropism, immune
 RT surveillance, and persistence of the organism.";
 RL J. Bacteriol. 183:5997-6008(2001).
 DR EMBL: AF304332; AAL14102.1; -;
 DR InterPro: IPR003506; Chlam_OMP6.
 DR Pfam: PF03504; Chlam_OMP6; 1.
 DR PRINTS: PR01356; CHLAMIDIOM6.
 FT NON_TER 511
 SQ SEQUENCE 511 AA: 54978 MW: 27555CB67A9214B CRC64;

Query Match 72.2%; Score 39; DB 2; Length 511;
 Best Local Similarity 70.0%; Pred. No. 26;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFVR 10
 Db 139 QLPCEAEFVR 148

RESULT 5
 093317 PRELIMINARY: PRT; 547 AA.
 AC 093317;
 DT 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Major outer membrane protein.
 CN OMPA.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-UM-12, AND UM-36;
 RX MEDLINE-21450826; PubMed-11567000;
 RA Millman K.L., Tavaré S., Dean D.;
 RT "Recombination in the ompA gene but Not the ompB gene of Chlamydia
 RT contributes to Serovar-Specific Differences in Tissue Tropism, Immune
 RT Surveillance, and Persistence of the Organism.";
 RL J. Bacteriol. 183:5997-6008(2001).
 DR EMBL: AF304330; AAL14100.1; -;
 DR EMBL: AF304332; AAL14099.1; -;
 DR InterPro: IPR003506; Chlam_OMP6.
 DR Pfam: PF03504; Chlam_OMP6; 1.
 DR PRINTS: PR01336; CHLAMIDIAOM6.
 SO SEQUENCE 547 AA; 5862 MW; 817BA3000FEA6A71 CRC64;

Query Match 72.2%; Score 39; DB 2; Length 547;
 Best Local Similarity 70.0%; Pred. No. 28;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLPCEAEFLR 10
 DB 139 QLPCEAEFVR 148

RESULT 6

ID 09LEB7 PRELIMINARY; PRT; 147 AA.
 AC 09LEB7;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Common plant regulatory factor 6.
 CN CPRF6.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; easterids II; Apiales; Apiaceae; Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, HAMBURGER SCHNITT;
 RX MEDLINE-21414626; PubMed-11523788;
 RA Ruegger A., Frohnmeyer H., Neeke C., Wellmer F., Kircher S.,
 RA Schaefer E., Hatter K.;
 RT "Isolation and characterization of four novel parsley proteins that
 RT interact with the transcriptional regulators CPRF1 and CPRF2.";
 RL Mol. Genet. Genomics 265:964-976(2001).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BRIP FAMILY.
 DR EMBL: AJ292744; CAC00657.1; -;
 DR InterPro: IPR004827; TF_bzip.
 DR Pfam: PF00170; bzip; 1.
 DR SMART: SMO0338; BRIZ; 1.
 DR PROSITE: PS00036; BRIP_BASIC; 1.
 KW DNA-binding; Nuclear protein.
 SO SEQUENCE 147 AA; 16719 MW; F03C83A5108C468D CRC64;

Query Match 70.4%; Score 38; DB 10; Length 147;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLPCEAE 7
 DB 130 QLPCEAQ 136

RESULT 7

ID 09PU23 PRELIMINARY; PRT; 265 AA.
 AC 09PU23;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Spou rRNA methylase family protein.
 CN TC0683.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MOPN / NIGG;
 RX MEDLINE-20150255; PubMed-10684935;
 RA Read T.D., Brunham R.C., Shen C., Gall S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Ueberback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwynn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL: AE002337; AAF73591.1; -;
 DR TIGR: TC0683; -;
 DR InterPro: IPR001537; Spou_methylase.
 DR Pfam: PF00588; Spou_methylase; 1.
 DR ProDom: PD001243; Spou_methylase; 1.
 KW Complete proteome.
 SO SEQUENCE 265 AA; 29855 MW; F5DEB8F5A3932915 CRC64;

Query Match 70.4%; Score 38; DB 16; Length 265;
 Best Local Similarity 66.7%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QLPCEAEFL 9
 DB 73 KLPCEAEFL 81

RESULT 8

ID 09A3L2 PRELIMINARY; PRT; 642 AA.
 AC 09A3L2;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Sensory box histidine kinase/response regulator.
 CN CC3191.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE-21173698; PubMed-11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potočka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Derooy R.T., Dodson R.J., Durkin A.S., Gwyn M.L., Halt D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Ueberback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
 CC REGULATORY COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.
 CC -1- SIMILARITY: TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.
 DR EMBL: AE005982; AAK25153.1; -;
 DR HSSP: P06657; 2CHF.

DR TIGR: CC3191; -
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR004358; Bact_sens_pr_C.
 DR InterPro: IPR003661; His_kinA.
 DR InterPro: IPR004359; His_kinA_sig.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000014; PAS_domain.
 DR InterPro: IPR001789; Response_reg.
 DR Pfam: PF02518; HATPase_C.1.
 DR Pfam: PF00072; Response_reg.1.
 DR Pfam: PF00512; signal.1.
 DR PRINTS: PRO0344; BCTRLSENSOR.
 DR ProDom: PD000039; Response_reg.1.
 DR SMART: SM00387; HATPase_C.1.
 DR SMART: SM00388; HSKA.1.
 DR SMART: SM00086; PAC.2.
 DR SMART: SM00091; PAS.1.
 DR SMART: SM00448; REC.1.
 DR TIGRPFAMs: TIGR00229; sensory_box.1.
 DR Kinase: Phosphorylation; Sensory transduction; Transferase;
 KW Complete proteome.
 SQ SEQUENCE 642 AA; 69795 MW; 83FDBC0FDD75478D CRC64;

Query Match 70.4%; Score 38; DB 16; Length 642;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CPABELL 10
 DB 340 CPABELL 346

RESULT 9
 O9W612 PRELIMINARY; PRT: 1339 AA.
 AC O9W612; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Vitellogenin.
 GN VTG.
 OS Pimephales promelas.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Pimephales.
 OC NCBI_TaxID=90988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Korte J.J., Kohl M.D., Jensen K.M., Pasha M.S., Parks L.G.,
 RA LeBlanc G.A., Ankley G.T.;
 RT *Fethead Minnow Vitellogenin: Complementary DNA Sequence and Messenger
 RT RNA and Protein Expression after 17(beta)-Estradiol Treatment.*;
 RL Environ. Toxicol. Chem. 19:972-981(2000).
 DR EMBL: AF130354; AAD23878.1; -
 DR InterPro: IPR001747; Lipid_transprt_N.
 DR Pfam: PF01347; Vitellogenin_N.1.
 SQ SEQUENCE 1339 AA; 146274 MW; C5DDF4EC4C3B1A2A CRC64;

Query Match 70.4%; Score 38; DB 13; Length 1339;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CPABELL 10
 DB 453 CPABELL 459

RESULT 10
 O9Y4B5 PRELIMINARY; PRT: 1353 AA.
 AC O9Y4B5; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE KIAA0802 protein (Fragment).
 GN KIAA0802.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=9087487; Pubmed=9872452;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
 RA Kotani H., Nomura N., Ohara O.;
 RT *Prediction of the coding sequences of unidentified human genes. XI.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.*;
 RL DNA Res. 5:277-286(1998).
 DR EMBL: AB018345; BAA34522.1; -
 FT NON_TER 1
 SQ SEQUENCE 1353 AA; 150386 MW; 27A38F75B6EBD03F CRC64;

Query Match 70.4%; Score 38; DB 4; Length 1353;
 Best Local Similarity 75.0%; Pred. No. 99;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 PCPABELL 10
 DB 247 PCPABELL 254

RESULT 11
 O90YN8 PRELIMINARY; PRT: 1362 AA.
 AC O90YN8; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Vitellogenin 1.
 OS Brachydanio rerio (zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Islinger M., Williams D., Voelkl A., Braunbeck T.;
 RT *Gene expression of estrogen-regulated genes associated with
 RT reproduction after exposure to 17alpha-ethinyltestradol in zebrafish
 RT (Danio rerio).*;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF406784; AAK94945.1; -
 DR InterPro: IPR001747; Lipid_transprt_N.
 DR Pfam: PF01347; Vitellogenin_N.1
 SQ SEQUENCE 1362 AA; 149547 MW; 35F8037F58F1DBDA CRC64;

Query Match 70.4%; Score 38; DB 13; Length 1362;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CPABELL 10
 DB 453 CPABELL 459

RESULT 12
 O916P4 PRELIMINARY; PRT: 284 AA.
 AC O916P4; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein PA0244.

GN PA0244.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 Gader R.L., Collier L., Tolentino E., Westbrock-Medman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Iarbig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL; AE004462; AAC03633.1;
 DR InterPro: IPR000205; NAD binding.
 DR InterPro: IPR002907; Shikimate_DH.
 DR Pfam: PF01488; Shikimate_DH; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 284 AA; 29691 MW; 1B5CF77FC7F473A CRC64;

Query Match 68.5%; Score 37; DB 16; Length 284;
 Best Local Similarity 77.8%; Pred. NO. 36;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPCPAELLR 10
 DB 211 MPVPAELLR 219

RESULT 13
 ID 094W74 PRELIMINARY; PRT; 348 AA.
 AC 094W74;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE NADH dehydrogenase subunit 2.
 OS Scatella histophorus (walking goby).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Gobioidae;
 OC Gobiidae; Scatellaos.
 ON NCBI_TaxID=166764;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SCA1ST;
 RA Thacker C.E.;
 RT "Molecular Phylogeny of the Gobioid Fishes."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIOUINONE -> NAD(+) + UBIOUINOL.
 DR EMBL; AF391490; AAL16591.1;
 DR InterPro: IPR001750; Oxidored_g1.
 DR Pfam: PF00361; Oxidored_g1; 1.
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 348 AA; 37652 MW; BE0272103A648FC3 CRC64;

Query Match 68.5%; Score 37; DB 8; Length 348;
 Best Local Similarity 66.7%; Pred. NO. 43;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 OLPCPAELL 9
 DB 144 OLPCPSOKL 152

RESULT 14

O9NFC7
 ID O9NFC7 PRELIMINARY; PRT; 405 AA.
 AC O9NFC7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CG18843 protein.
 GN CG18843.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mates B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA GIBBS R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003423; AAG32398.1;
 DR FlyBase; FBgn0042163; CG18843.
 SQ SEQUENCE 405 AA; 45063 MW; DCCA90527469769B CRC64;

Query Match 68.5%; Score 37; DB 5; Length 405;
 Best Local Similarity 75.0%; Pred. NO. 50;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPCPAELL 9
 DB 140 LPCPVDL 147

RESULT 15
 ID 094909 PRELIMINARY; PRT; 989 AA.
 AC 094909;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE KIAA0819 protein (Fragment).
 GN KIAA0819.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE-99156230; PubMed-10048485;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hiroseawa M.,
 Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT *Prediction of the coding sequences of unidentified human genes.
 RT XII. The complete sequences of 100 new cDNA clones from brain which
 code for large proteins in vitro.";
 RL DNA Res. 5:355-364(1998).
 DR EMBL; AB020626; BAA74842.1; "
 DR InterPro: IPR002965; P_Rich_extensn.
 DR PRINTS: PR01217; PRICHEXTENSN.
 FT NON_TER 1
 SQ SEQUENCE 989 AA; 109888 MW; F8845C9CB90D7671 CRC64;
 QY 1 QLPCPAE 7
 :|||||
 Db 105 RLPCPAE 111
 Query Match 68.5%; Score 37; DB 4; Length 989;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Search completed: July 9, 2003, 12:02:01
 Job time : 38.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:55:35 ; Search time 22 Seconds

(without alignments)
30.284 Million cell updates/sec

Title: US-09-851-058-2

Perfect score: 30

Sequence: 1 YSQCR 5

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.GeneSeq_101002:*

1: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*

2: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*

3: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*

4: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*

5: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*

6: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*

7: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*

8: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*

9: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*

10: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*

11: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*

12: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*

13: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*

14: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*

15: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*

16: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*

17: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*

18: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*

19: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*

20: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*

21: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*

22: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*

23: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	87	22	AA008211
2	30	100.0	243	21	AAV28425
3	27	90.0	39	21	AA612153
4	27	90.0	52	22	AA676769
5	27	90.0	92	22	AB616866
6	27	90.0	124	21	AA640523
7	27	90.0	124	23	ABP31173
8	27	90.0	134	19	AAW59178
9	27	90.0	153	22	AB609610
10	27	90.0	155	19	AAW59177

11	27	90.0	155	20	AAW94272
12	27	90.0	155	20	AAW94273
13	27	90.0	187	22	AB613128
14	27	90.0	320	22	AB658567
15	27	90.0	424	21	AA651387
16	27	90.0	585	22	AB658555
17	27	90.0	755	22	AB658201
18	27	90.0	1693	12	AA646184
19	27	90.0	1693	15	AA651264
20	27	90.0	1693	17	AA631813
21	27	90.0	1693	19	AA681519
22	27	90.0	1693	19	AA680196
23	27	90.0	1693	19	AA676368
24	27	90.0	1693	19	AA671209
25	27	90.0	1693	21	AA624119
26	27	90.0	1693	22	AA625222
27	27	90.0	1698	20	AA631381
28	27	90.0	1707	22	AA678760
29	27	90.0	1708	20	AA631384
30	27	90.0	1708	20	AA63408
31	27	90.0	2228	22	AB609042
32	26	86.7	32	12	AA63252
33	26	86.7	46	17	AA60658
34	26	86.7	51	17	AA60656
35	26	86.7	53	18	AA601850
36	26	86.7	58	18	AA601845
37	26	86.7	58	18	AA601844
38	26	86.7	61	19	AA621451
39	26	86.7	63	22	AB627573
40	26	86.7	69	22	AA645102
41	26	86.7	71	22	AA644863
42	26	86.7	75	22	AA64472
43	26	86.7	78	22	AA652439
44	26	86.7	98	17	AA600663
45	26	86.7	103	22	AB644333

ALIGNMENTS

RESULT 1	AA008211
ID	AA008211 standard; Protein: 87 AA.
AC	AA008211;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 22103.
XX	
KW	Human: cytokine; cell proliferation; gene therapy;
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW	tissue growth factor; immunomodulatory; cancer; leukemia;
KW	nervous system disorders; arthritis; inflammation.
OS	Homo sapiens.
XX	
PN	WO200164835-A2.
XX	
PD	07-SEP-2001.
XX	
PF	26-FEB-2001; 2001WO-US04927.
XX	
PR	28-FEB-2000; 2000US-0515126.
XX	
PA	18-MAY-2000; 2000US-0577409.
XX	
PI	(HYSE-) HYSEQ INC.
XX	
DR	Tang YT, Liu C, Drmanac RT;
XX	
DR	WPI: 2001-514838/56.
XX	
DR	N-PSDB; AA188142.
XX	

Rat-derived eosino
Rat eosinophils-de
Novel human diagno
Drosophila melanog
Arabidopsis thalia
Drosophila melanog
Drosophila melanog
Protein encoded by
HEV strain protein
Hepatitis E virus
Hepatitis E virus
Hepatitis E virus
Protein encoded by
Hepatitis E virus
Hepatitis E virus
HEV-Burma strain v
HEV-051 ORF1 prote
Hepatitis E virus
HEV-052 ORF1 prote
Swine HEV ORF 1 pr
Novel human diagno
Carbohydrate bindi
Mucin #1 of secre
Antileukoprotease
Antileukoprotease
Antileukoprotease
Human neuroendocr
Novel human diagno
Propionibacterium
Propionibacterium
Propionibacterium
Mucin #1 of secre
Peptide #11839 enc

PT Isolated nucleic acids and polypeptides, useful for preventing
PT disorders and treating e.g. leukemia, inflammation and immune
PT disorders.
PS
XX
XX Claim 20: SEQ ID NO 22103; 1399pp + Sequence Listing: English.
CC
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, hematopoietic regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcc_sequences.
XX
SQ Sequence 87 AA:

Query Match 100.0%; Score 30; DB 22; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSOCR 5
DB 58 YSOCR 62
IIIIII

RESULT 2
AAV28425
ID AAV28425 standard; Protein: 243 AA.
AC AAV28425;
XX
XX 15-FEB-2000 (first entry)
DE
XX Rice branched chain amino acid transaminase amino acid sequence.
XX
XX Corn; soybean; wheat; rice: dihydroxyacid dehydratase; leuc; leud;
XX branched chain amino acid transferase; biosynthetic enzyme; antibody;
XX 3-isopropylmalate dehydratase.
XX
XX Oryza sativa.
XX
XX WO9921880-A2.
XX
XX 06-MAY-1999.
XX
XX 20-OCT-1998; 98WO-US22081.
XX
XX 28-OCT-1997; 97US-0063423.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
XX Falco SC, Cahoon RE, Hiltz WD, Kinney AJ, Rafalski JA;
XX
XX WPI: 2000-022904/02.
XX
XX N-PSDB: AAX89451.
XX
XX Nucleic acid fragments encoding branched chain amino acid biosynthetic
XX enzymes
XX
XX Claim 6; Fig 4; 102pp; English.
XX
XX AAV28418-Y28431 are fragments of corn, soybean, wheat and rice branched
XX chain amino acid transaminase, amino acid sequences. Sequences
XX AAX89442-X89465 are nucleic acid fragments that encode all or a
XX substantial portion of dihydroxyacid dehydratase, a branched chain amino
XX acid transferase, a leuc or a leud subunit of 3-isopropylmalate

CC dehydratase from wheat, corn, soybean or rice. These enzymes are involved
CC in biosynthesis and utilization of branched-chain amino acids. The
CC nucleic acid sequences can be used to alter the level of expression of a
CC branched chain amino acid biosynthetic enzyme in a host cell. They can
CC also be used to obtain a nucleic acid fragment encoding a branched chain
CC amino acid biosynthetic enzyme, and to identify inhibitors of a branched
CC chain amino acid biosynthetic enzyme. The encoded proteins may be used
CC to prepare antibodies for detecting the proteins in situ in cells, or in
CC vitro in cell extracts.
XX
XX
SQ Sequence 243 AA:

Query Match 100.0%; Score 30; DB 21; Length 243;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSOCR 5
DB 68 YSOCR 72
IIIIII

RESULT 3
AAG12153
ID AAG12153 standard; Protein: 39 AA.
AC AAG12153;
XX
XX 17-OCT-2000 (first entry)
DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 11153.
XX
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX
XX 05-MAR-1999; 99US-0123180.
XX
XX 09-MAR-1999; 99US-0123548.
XX
XX 23-MAR-1999; 99US-0125788.
XX
XX 25-MAR-1999; 99US-0126264.
XX
XX 29-MAR-1999; 99US-0126785.
XX
XX 01-APR-1999; 99US-0127462.
XX
XX 06-APR-1999; 99US-0128234.
XX
XX 08-APR-1999; 99US-0128714.
XX
XX 16-APR-1999; 99US-0129845.
XX
XX 19-APR-1999; 99US-0130077.
XX
XX 21-APR-1999; 99US-0130449.
XX
XX 23-APR-1999; 99US-0130510.
XX
XX 28-APR-1999; 99US-0130891.
XX
XX 30-APR-1999; 99US-0131449.
XX
XX 30-APR-1999; 99US-0132048.
XX
XX 04-MAY-1999; 99US-0132407.
XX
XX 05-MAY-1999; 99US-0132484.
XX
XX 06-MAY-1999; 99US-0132485.
XX
XX 06-MAY-1999; 99US-0132486.
XX
XX 07-MAY-1999; 99US-0132487.
XX
XX 11-MAY-1999; 99US-0132863.
XX
XX 14-MAY-1999; 99US-0134256.
XX
XX 14-MAY-1999; 99US-0134218.
XX
XX 14-MAY-1999; 99US-0134219.
XX
XX 14-MAY-1999; 99US-0134221.
XX
XX 14-MAY-1999; 99US-0134370.
XX
XX 18-MAY-1999; 99US-0134768.
XX
XX 20-MAY-1999; 99US-0134941.
XX
XX 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145219.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 20-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153738.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159299.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match
 Best Local Similarity 80.0%; Score 27; DB 21; Length 39;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
 1:111
 Db 14 YTOCR 18

RESULT 4
 AAG76769
 ID AAG76769 standard; Protein; 52 AA.
 AC AAG76769;
 XX
 DT 03-SEP-2001 (first entry)
 DE Human colon cancer antigen protein SEQ ID NO:7533.
 XX
 DE Human colon cancer antigen protein SEQ ID NO:7533.
 XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 DR WPI; 2001-235357/24.
 DR N-PSDB; AAG36174.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11; Page 8951-8952; 9803pp; English.
 CC AAG72943 to AAG73195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAG73196 to AAG7204
 CC and AAG77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SO Sequence 52 AA;
 Query Match 90.0%; Score 27; DB 22; Length 52;

Best Local Similarity 80.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
 1:111
 Db 29 YNOCR 33

RESULT 5
 ABG16866
 ID ABG16866 standard; Protein; 92 AA.
 AC ABG16866;
 XX
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #16857.
 XX
 DE Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS81053.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 47225; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 92 AA;
 Query Match 90.0%; Score 27; DB 22; Length 92;
 Best Local Similarity 80.0%; Pred. No. 5.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
|||||
Db 7 YSOCR 11

RESULT 6
AAB40523
ID AAB40523 standard; Protein: 124 AA.

AC AAB40523;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF287 polypeptide sequence SEQ ID NO:574.

XX Human: open reading frame; ORFX: detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; noctropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.

OS Homo sapiens.

XX MO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000MO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC74732.

XX Claim 11; Page 719; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; noctropic; neuroprotective;
CC anticonvulsant; osteopathic; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotrophic;
CC immunosuppressive; antineoplastic; antirheumatic; antihypertoid;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antihypertoid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus

CC erythematous, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 124 AA;

QY Query Match 90.0%; Score 27; DB 21; Length 124;

Best Local Similarity 80.0%; Pred. No. 7e+02; Mismatches 0; Gaps 0;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 110 YAOOR 114

RESULT 7
ABP31173
ID ABP31173 standard; Protein: 124 AA.

XX ABP31173;

XX 08-JUL-2002 (first entry)

DE Human ORF146 protein, SEQ ID NO:292.

XX Human: ORF: open reading frame; ORFX: drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotrophic; antipsoriatic; antidiabetic; cytostatic; noctropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antihypertoid; antinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX Homo sapiens.

XX MO200190366-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001MO-US17076.

XX 24-MAY-2000; 2000US-206690P.

XX (CURA-) CURAGEN CORP.

XX Leach MD, Shimkets RA;

XX WPI; 2002-106200/14.

XX N-PSDB; ABN75199.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation

XX Claim 10; Page 342; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79387 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and

CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, hematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration, to
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.
XX

SQ Sequence 124 AA:

Query Match Best Local Similarity 90.0%; Score 27; DB 23; Length 124;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSQR 5
11111
Db 110 YACR 114

RESULT 8
AAM59178
ID AAM59178 standard; Protein; 134 AA.
XX
AC AAM59178;

DT 13-AUG-1998 (first entry)

DE Rat eosinophil cationic protein #2.

XX
KM Eosinophil cationic protein; ECP; cytotoxic; recombinant protein;
treatment; disease; rat.

XX
OS Rattus sp.

XX
PN JP1011777-A.

XX
PD 12-MAY-1998.

XX
PE 25-OCT-1996; 96JP-0284311.

XX
PR 25-OCT-1996; 96JP-0284311.

XX
PA (NIH) JAPAN ENERGY CORP.

XX
DR WPI; 1998-325891/29.

XX
PT New DNA coding rat-derived eosinophil cationic protein - used, e.g.
to produce recombinant protein for treatment of diseases

XX
PS Claim 1; Page 12; 14pp; Japanese.

XX
CC This sequence is a novel rat-derived eosinophil cationic protein (ECP)
CC which exhibits cytotoxic activity. The DNA encoding the ECP protein
CC can be used for the production of recombinant protein which is useful

CC for the treatment of diseases.

XX
SQ Sequence 134 AA;

Query Match Best Local Similarity 90.0%; Score 27; DB 19; Length 134;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSQR 5
11111
Db 94 YOCR 98

RESULT 9
ABG09610
ID ABG09610 standard; Protein; 153 AA.
XX
AC ABG09610;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9601.

XX
KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

XX
OS Homo sapiens.

XX
PN WO200175067-A2.

XX
PD 11-OCT-2001.

XX
PE 30-MAR-2001; 2001WO-US08631.

XX
PR 31-MAR-2000; 2000US-0540217.

XX
PX 23-AUG-2000; 2000US-0649167.

XX
PA (HSE-) HYSEQ INC.

XX
PI Drmanac RT, Liu C, Tang YH;

XX
DR WPI; 2001-639362/73.

XX
DR N-PSDB; AAS73797.

XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX
PS Claim 20; SEQ ID NO 39969; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

PT Preparing recombinant protein
 XX
 PS Disclosure; Fig 4; 18pp; Japanese.
 CC This represents a rat eosinophils-derived eosinophil cationic protein (ECP). The DNA encoding a rat-derived eosinophil-derived neurotoxin (EDN) or its precursor peptide is useful for the preparation of a recombinant protein. The base sequence of a precursor peptide of rat eosinophils-derived ECP, and the base sequence of the precursor peptide of EDN can be determined by collecting the ECP cDNA and EDN cDNA from eosinophils respectively.
 CC Sequence 155 AA;
 SO
 Query Match 90.0%; Score 27; DB 20; Length 155;
 Best Local Similarity 80.0%; Pred. No. 8.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YSQCR 5
 1:111
 DB 115 YSQCR 119
 RESULT 13
 ABG13128
 ID ABG13128 standard; Protein; 187 AA.
 XX
 AC ABG13128;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #13119.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS77315.
 PT
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 PS Claim 20; SEQ ID NO 43487; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SO Sequence 187 AA;
 Query Match 90.0%; Score 27; DB 22; Length 187;
 Best Local Similarity 80.0%; Pred. No. 9.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YSQCR 5
 1:111
 DB 140 YSQCR 144
 RESULT 14
 ABB58567
 ID ABB58567 standard; Protein; 320 AA.
 XX
 AC ABB58567;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 2493.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 PI
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL02670.
 PT
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Disclosure; SEQ ID NO 2493; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB12072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SO Sequence 320 AA;
 Query Match 90.0%; Score 27; DB 22; Length 320;

Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YSOCR 5
Db 229 YAOCR 233

RESULT 15
AAG51387
ID AAG51387 standard; Protein: 424 AA.

XX AAC51387;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65214.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

PD 06-SEP-2000.

FE 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128273.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137502.

PR 04-JUN-1999; 99US-0137724.

PR 07-JUN-1999; 99US-0138094.

PR 08-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142154.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 21-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 03-AUG-1999; 99US-0146389.

PR 04-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 05-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 06-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 09-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 10-AUG-1999; 99US-0147933.

PR 11-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

Search completed: July 9, 2003, 12:00:06
Job time : 23 secs

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161892.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 90.0%; Score 27; DB 21; Length 424;
Pred. No. 2e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOQR 5
11:11
Db 10 YSECR 14

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:58:36 ; Search time 11.25 Seconds
(without alignments)
13.077 Million cell updates/sec

Title: US-09-851-058-2

Perfect score: 30

Sequence: 1 YSCOR 5

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued_Patents_AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/prodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/prodata/1/1aa/5A_COMB.pep.*

4: /cgn2_6/prodata/1/1aa/5B_COMB.pep.*

5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/prodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	243	4	US-09-173-300-22
2	27	90.0	1693	3	US-08-840-316-1
3	27	90.0	1693	3	US-08-478-507-7
4	27	90.0	1693	4	US-08-809-523-1
5	27	90.0	1693	4	US-09-128-275A-7
6	27	90.0	1693	4	US-08-471-971-1
7	27	90.0	1693	4	US-09-553-427-7
8	27	90.0	1693	4	US-09-402-776-1
9	27	90.0	1693	5	PCT-US93-08849A-1
10	27	90.0	1693	5	PCT-US93-08849A-1
11	27	90.0	1708	4	US-09-462-606-2
12	26	86.7	32	1	US-08-361-820-10
13	26	86.7	32	1	US-08-479-939-10
14	26	86.7	32	1	US-08-483-432-10
15	26	86.7	47	2	US-08-464-182A-27
16	26	86.7	47	2	US-08-693-274A-16
17	26	86.7	47	2	US-08-406-271-27
18	26	86.7	108	3	US-08-959-212-10
19	26	86.7	155	1	US-08-150-203A-9
20	26	86.7	155	1	US-08-454-730-9
21	26	86.7	155	4	US-08-949-788-9
22	26	86.7	216	4	US-09-556-877-20
23	26	86.7	216	4	US-09-288-594A-20
24	26	86.7	216	4	US-09-620-412C-20
25	26	86.7	291	2	US-08-560-098A-55
26	26	86.7	295	4	US-09-799-345-2
27	26	86.7	349	3	US-08-651-136C-10

28	26	86.7	349	4	US-09-229-911A-10	Sequence 10, Appl
29	26	86.7	377	4	US-09-410-464-8	Sequence 8, Appl
30	26	86.7	424	2	US-08-592-214A-16	Sequence 16, Appl
31	26	86.7	424	3	US-08-559-188-16	Sequence 16, Appl
32	26	86.7	424	3	US-08-555-227-16	Sequence 16, Appl
33	26	86.7	424	3	US-08-655-241-16	Sequence 16, Appl
34	26	86.7	424	3	US-09-149-976-16	Sequence 16, Appl
35	26	86.7	424	4	US-09-398-326-16	Sequence 16, Appl
36	26	86.7	428	3	US-09-118-319-5	Sequence 5, Appl
37	26	86.7	432	3	US-09-118-319-9	Sequence 2, Appl
38	26	86.7	461	3	US-09-118-319-9	Sequence 2, Appl
39	26	86.7	514	1	US-08-361-920-21	Sequence 21, Appl
40	26	86.7	514	1	US-08-479-939-21	Sequence 21, Appl
41	26	86.7	514	1	US-08-483-432-21	Sequence 21, Appl
42	26	86.7	575	4	US-08-924-345-3	Sequence 3, Appl
43	26	86.7	577	4	US-08-484-105-24	Sequence 24, Appl
44	26	86.7	577	1	US-08-484-106-24	Sequence 24, Appl
45	26	86.7	608	4	US-09-134-001C-3395	Sequence 3395, Ap

ALIGNMENTS

RESULT 1
US-09-173-300-22
Sequence 22, Application US/09173300
Patent No. 6451581
GENERAL INFORMATION:
APPLICANT: Falco, Saverio Carl
APPLICANT: Hiltz, William D.
APPLICANT: Kinney, Anthony J.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, J. Antoni
TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
FILE REFERENCE: BB-1126
CURRENT APPLICATION NUMBER: US/09/173,300
CURRENT FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: 60/063,423
EARLIER FILING DATE: 1997 October 28
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Microsoft Word Version 7.0A
SEQ ID NO 22
LENGTH: 243
TYPE: PRT
ORGANISM: Glycine max
US-09-173-300-22

Query Match 100.0%; Score 30; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 YSCOR 5
Db 68 YSCOR 72

RESULT 2
US-08-840-316-1
Sequence 1, Application US/08840316
Patent No. 6054567
GENERAL INFORMATION:
APPLICANT: Emerson, Suzanne U., Purcell, Robert H.,
APPLICANT: Tsarev, Sergei, A., and Robinson, Robin A.
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain of Hepatitis E And Their
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA

ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840.316
FILING DATE: 11-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4255
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 AMINO ACID RESIDUES
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-840-316-1

Query Match 90.0%; Score 27; DB 3; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5 1:111
DB 431 YACR 435

RESULT 3
US-08-478-507-7
Sequence 7, Application US/08478507
Patent No. 6120988
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R
APPLICANT: Yarbough, Patrice O
APPLICANT: Birdley, Daniel W
APPLICANT: Krawczynski, Krzysztof Z
APPLICANT: Fry, Kirk E
TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
NUMBER OF SEQUENCES: No. 6120988-A/No. 6120988-B Hepatitis Viral Agent
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478.507
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279.823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681.078
FILING DATE: 05-APR-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-507-7

Query Match 90.0%; Score 27; DB 3; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5 1:111
DB 431 YACR 435

RESULT 4
US-08-809-523-1
Sequence 1, Application US/08809523
Patent No. 6207416
GENERAL INFORMATION:
APPLICANT: Tsarev, Sergei A., Emerson.
APPLICANT: Suzanne U., Purcell, Robert H.
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSER: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,523
FILING DATE: 28-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13102
FILING DATE: 03-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/316,765
FILING DATE: 03-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/947,263

FILED DATE: 18-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4032054
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 AMINO ACID RESIDUES
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-809-523-1

Query Match 90.0%; Score 27; DB 4; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YSQR 5
1:1111
Db 431 YACR 435

RESULT 5
US-09-128-275A-7
Sequence 7, Application US/09128275A
Patent No. 6229005
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R
APPLICANT: Yarbough, Patrice O
APPLICANT: Bradley, Daniel W
APPLICANT: Krawczynski, Krzysztof Z
APPLICANT: Tam, Albert
TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128, 275A
FILING DATE: 03-AUG-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279, 823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681, 078
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505, 888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420, 921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367, 486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336, 672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/208, 997
FILED DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Pelthor, Joanne R.
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 4600-0183, 24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-128-275A-7

Query Match 90.0%; Score 27; DB 4; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YSQR 5
1:1111
Db 431 YACR 435

RESULT 6
US-08-471-971-1
Sequence 1, Application US/08471971
Patent No. 6287759
GENERAL INFORMATION:
APPLICANT: Tsarev, Sergei A., Emerson,
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & PINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471, 971
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/316, 765
FILING DATE: 03-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US07/947, 263
FILING DATE: 18-SEP-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4032052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 AMINO ACID RESIDUES
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN

TOPOLOGY: UNKNOWN
US-08-471-971-1

Query Match 90.0%; Score 27; DB 4; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
1:111
DB 431 YAOCR 435

RESULT 7
US-09-553-427-7

Sequence 7, Application US/09553427
Patent No. 6379891

GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R

APPLICANT: Yarbough, Patricia O

APPLICANT: Bradley, Daniel W

APPLICANT: Krawczynski, Krzysztof Z

APPLICANT: Fry, Kirk E

TITLE OF INVENTION: DNA Sequences of Enterically Transmitted

TITLE OF INVENTION: No. 6379891-A/No. 6379891-B Hepatitis Viral Agent

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/553,427

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/478,507

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/279,823

FILING DATE: 25-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/681,078

FILING DATE: 05-APR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/505,888

FILING DATE: 05-APR-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/420,921

FILING DATE: 13-OCT-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/367,486

FILING DATE: 16-JUN-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/336,672

FILING DATE: 11-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/208,997

FILING DATE: 17-JUN-1988

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 4600-0183, 22

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960

INFORMATION FOR SEQ ID NO: 7:

LENGTH: 1693 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-553-427-7

Query Match 90.0%; Score 27; DB 4; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
1:111
DB 431 YAOCR 435

RESULT 8
US-09-402-776-1

Sequence 1, Application US/09402776

Patent No. 6458562

GENERAL INFORMATION:

APPLICANT: Emerson, Suzanne U., Purcell, Robert H.,

APPLICANT: Tsarev, Sergei, A., and Robinson, Robin A.

TITLE OF INVENTION: Recombinant Proteins of

TITLE OF INVENTION: A Pakistani Strain of Hepatitis E And Their

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/402,776

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/840,316

FILING DATE: 11-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Richard W. Bork

REGISTRATION NUMBER: 36,459

REFERENCE/DOCKET NUMBER: 2026-4255

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1693 AMINO ACID RESIDUES

TYPE: AMINO ACID

STRANDEDNESS: UNKNOWN

TOPOLOGY: UNKNOWN

US-09-402-776-1

Query Match 90.0%; Score 27; DB 4; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
1:111
DB 431 YAOCR 435

RESULT 9

PCT-US93-08849A-1

Sequence 1, Application PC/TUS9308849A

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08849A
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US07/947,263
FILING DATE: 18-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4032 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 AMINO ACID RESIDUES
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
PCT-US93-08849A-1

Query Match 90.0%; Score 27; DB 5; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
|:|:|:|
Db 431 YACR 435

RESULT 10
PCT-US93-08849-1
Sequence 1, Application PC/TUS9308849
GENERAL INFORMATION:
APPLICANT: Tsarev, Sergei A., Emerson,
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08849
FILING DATE: 17-SEP-1993
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/947,263
FILING DATE: 18-SEP-1992
NAME:
ATTORNEY/AGENT INFORMATION:
NAME: Bork, Richard, W.
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 amino acid residues
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
PCT-US93-08849-1

Query Match 90.0%; Score 27; DB 5; Length 1693;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
|:|:|:|
Db 431 YACR 435

RESULT 11
US-09-462-606-2
Sequence 2, Application US/09462606
Patent No. 6432408
GENERAL INFORMATION:
APPLICANT: MENG, XIANG-JIN
APPLICANT: Emerson, Suzanne U.
TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
FILE REFERENCE: 20264267US1
CURRENT APPLICATION NUMBER: US/09/462,606
PRIOR FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: US 60/053069
PRIOR FILING DATE: 1997-07-18
PRIOR APPLICATION NUMBER: PCT/US98/14665
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1708
TYPE: PRT
ORGANISM: Hepatitis E virus
US-09-462-606-2

Query Match 90.0%; Score 27; DB 4; Length 1708;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
|:|:|:|
Db 431 YACR 435

RESULT 12
US-08-361-920-10
Sequence 10, Application US/08361920
Patent No. 5457046
GENERAL INFORMATION:
APPLICANT: woelike, Helle F.
APPLICANT: Hagen, Frederick
APPLICANT: Hjort, Carsten M.
APPLICANT: Sven, Hastrup
TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
TITLE OF INVENTION: or Hemicellulose
NUMBER OF SEQUENCES: 85

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 54570460 No. 5457046disk of No. 5457046th America, Inc.
;; STREET: 405 Lexington Avenue, 62nd Floor
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States of America
;; ZIP: 10174-6201
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/361,920
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/940,860
;; FILING DATE: 28-OCT-1992
;; APPLICATION NUMBER: DK 1158/90
;; FILING DATE: 09-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/DK91/00124
;; FILING DATE: 08-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lambiris, Elias J.
;; REGISTRATION NUMBER: 33,728
;; REFERENCE/DOCKET NUMBER: 3435.204-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-867-0298
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 32 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHEICAL: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Fusarium oxysporum
;;
;; US-08-361-920-10
;;
Query Match 86.7%; Score 26; DB 1; Length 32;
Best Local Similarity 80.0%; Pred. No. 83;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSQC 5
DB 28 YSQC 32
;;
RESULT 13
US-08-479-939-10
;; Sequence 10, Application US/08479939
;; Patent No. 5686593
;; GENERAL INFORMATION:
;; APPLICANT: Woeldike, Helle F.
;; APPLICANT: Hagen, Frederick
;; APPLICANT: Hjort, Carsten M.
;; APPLICANT: Sven, Hasstrup
;; TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
;; NUMBER OF SEQUENCES: 85
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 56865930 No. 5686593disk of No. 5686593th America, Inc.
;; STREET: 405 Lexington Avenue, 62nd Floor
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States of America
;; ZIP: 10174-6201
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/479,939
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/361,920
;; FILING DATE: 22-DEC-1994
;; APPLICATION NUMBER: US 07/940,860
;; FILING DATE: 28-OCT-1992
;; APPLICATION NUMBER: DK 1158/90
;; FILING DATE: 09-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/DK91/00124
;; FILING DATE: 08-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lambiris, Elias J.
;; REGISTRATION NUMBER: 33,728
;; REFERENCE/DOCKET NUMBER: 3435.204-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-867-0298
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 32 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHEICAL: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Fusarium oxysporum
;;
;; US-08-479-939-10
;;
Query Match 86.7%; Score 26; DB 1; Length 32;
Best Local Similarity 80.0%; Pred. No. 83;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSQC 5
DB 28 YSQC 32
;;
RESULT 14
US-08-483-432-10
;; Sequence 10, Application US/08483432
;; Patent No. 5763254
;; GENERAL INFORMATION:
;; APPLICANT: Woeldike, Helle F.
;; APPLICANT: Hagen, Frederick
;; APPLICANT: Hjort, Carsten M.
;; APPLICANT: Sven, Hasstrup
;; TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
;; NUMBER OF SEQUENCES: 85
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 57632540 No. 5763254disk of No. 5763254th America, Inc.
;; STREET: 405 Lexington Avenue, 62nd Floor
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States of America
;; ZIP: 10174-6201
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/483,432
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/361,920
FILING DATE: 07-09-1992
APPLICATION NUMBER: US 07/940,860
FILING DATE: 28-OCT-1992
APPLICATION NUMBER: DK 1158/90
FILING DATE: 09-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00124
FILING DATE: 08-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3435, 204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-867-0298
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Fusarium oxysporum
US-08-483-432-10

Query Match 86.7%; Score 26; DB 1; Length 32;
Best Local Similarity 80.0%; Pred. No. 83;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
11111
DB 28 YSOGR 32

RESULT 15
US-08-464-182A-27
Sequence 27, Application US/08464182A
Patent No. 5849885
GENERAL INFORMATION:
APPLICANT: Nuijens, Jan H.
APPLICANT: Van Veen, Harry H.
TITLE OF INVENTION: Isolation of Lactoferrin from milk
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,182A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/406,271
FILING DATE: 09-MAR-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,321
FILING DATE: 16-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 016994-004920US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 326-2400
TELEFAX: (650) 326-2422
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..47
OTHER INFORMATION: /note="PLF"
US-08-464-182A-27

Query Match 86.7%; Score 26; DB 2; Length 47;
Best Local Similarity 80.0%; Pred. No. 1,2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
11111
DB 16 YSKCR 20

Search completed: July 9, 2003, 12:03:36
Job time: 13.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:02:11 : Search time 20.25 Seconds
(Without alignments)
28.746 Million cell updates/sec

Title: US-09-851-058-2
Perfect score: 30
Sequence: 1 YSOCR 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 445758 seqs, 116419773 residues
Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubppaa/PC7_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubppaa/PC7US_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	30	100.0	243	12 US-10-027-450-22	Sequence 22, Appl
2	27	90.0	52	9 US-10-106-698-7543	Sequence 7543, Ap
3	27	90.0	1693	9 US-09-851-410-7	Sequence 7, Appl
4	27	90.0	1698	9 US-09-468-147-91	Sequence 91, Appl
5	27	90.0	1708	9 US-09-468-147-166	Sequence 166, App
6	26	86.7	103	10 US-09-864-761-42491	Sequence 42491, A
7	26	86.7	114	10 US-09-925-300-1551	Sequence 1551, Ap
8	26	86.7	199	10 US-09-734-569-130	Sequence 130, App
9	26	86.7	216	10 US-09-841-132-20	Sequence 20, Appl
10	26	86.7	253	10 US-09-925-300-1601	Sequence 1601, Ap
11	26	86.7	289	9 US-09-738-626-5525	Sequence 5525, Ap
12	26	86.7	295	10 US-09-962-276-2	Sequence 2, Appl
13	26	86.7	349	9 US-10-007-521-10	Sequence 10, Appl
14	26	86.7	356	9 US-10-156-761-8811	Sequence 8811, Ap
15	26	86.7	377	9 US-10-104-580-8	Sequence 8, Appl
16	26	86.7	381	9 US-10-052-092-29	Sequence 29, Appl
17	26	86.7	420	10 US-09-853-386-129	Sequence 129, App
18	26	86.7	424	9 US-09-853-450-16	Sequence 16, Appl
19	26	86.7	577	9 US-09-893-519A-61	Sequence 61, Appl

20	26	86.7	591	10 US-09-815-242-5662	Sequence 5662, Ap
21	26	86.7	1000	9 US-09-823-187-6	Sequence 6, Appl
22	25	83.3	25	9 US-09-983-802-152	Sequence 152, App
23	25	83.3	30	10 US-09-917-340-49	Sequence 49, Appl
24	25	83.3	31	9 US-09-916-494A-2	Sequence 2, Appl
25	25	83.3	33	9 US-09-916-494A-6	Sequence 6, Appl
26	25	83.3	34	7 US-08-841-636A-37	Sequence 37, Appl
27	25	83.3	34	12 US-10-001-843-205	Sequence 205, App
28	25	83.3	39	9 US-09-916-494A-4	Sequence 4, Appl
29	25	83.3	40	9 US-10-125-258-103	Sequence 103, App
30	25	83.3	40	9 US-10-125-258-107	Sequence 107, App
31	25	83.3	40	9 US-10-125-258-111	Sequence 111, App
32	25	83.3	46	9 US-10-073-961-283	Sequence 283, App
33	25	83.3	46	10 US-09-764-887-283	Sequence 283, App
34	25	83.3	52	9 US-09-983-802-301	Sequence 301, App
35	25	83.3	64	9 US-10-125-258-101	Sequence 101, App
36	25	83.3	64	9 US-10-125-258-105	Sequence 105, App
37	25	83.3	64	9 US-10-125-258-109	Sequence 109, App
38	25	83.3	71	9 US-10-083-357-763	Sequence 763, App
39	25	83.3	73	10 US-09-864-761-40832	Sequence 40832, A
40	25	83.3	102	8 US-08-424-550B-44	Sequence 44, Appl
41	25	83.3	141	9 US-10-101-464A-91	Sequence 91, Appl
42	25	83.3	144	9 US-10-156-761-11527	Sequence 11527, A
43	25	83.3	145	10 US-09-893-737-88	Sequence 88, Appl
44	25	83.3	156	9 US-09-907-969-586	Sequence 586, App
45	25	83.3	156	9 US-10-198-053-586	Sequence 586, App

ALIGNMENTS

RESULT 1
US-10-027-450-22
Sequence 22, Application US/10027450
Patent No. US20020102715A1
GENERAL INFORMATION:
APPLICANT: Falco, Saverio Carl
APPLICANT: Hite, William D.
APPLICANT: Kinney, Anthony J.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, J. Antoni
TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMS
FILE REFERENCE: BR-1126
CURRENT APPLICATION NUMBER: US/10/027,450
CURRENT FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/063,423
PRIOR FILING DATE: 1997 October 28
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Microsoft Word Version 7.0A
SEQ ID NO 22
LENGTH: 243
TYPE: PRT
ORGANISM: Glycine max
US-10-027-450-22

Query Match 100.0%; Score 30; DB 12; Length 243;
Best Local Similarity 100.0%; Pred. No. 1;le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
|||||
DB 68 YSOCR 72

RESULT 2
US-10-106-698-7543
Sequence 7543, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698

;; CURRENT FILING DATE: 2002-03-27
;; PRIOR APPLICATION NUMBER: PCT/US00/26524
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: US 60/157,137
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/163,280
;; PRIOR FILING DATE: 1999-11-03
;; NUMBER OF SEQ ID NOS: 8564
;; SOFTWARE: PatentIn Ver. 3.0
;; SEQ ID NO 7543
;; LENGTH: 52
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-106-698-7543

Query Match 90.0%; Score 27; DB 9; Length 52;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
1:111
DB 29 YNOCR 33

RESULT 3

US-09-851-410-7
Sequence 7, Application US/09851410
Publication No. US20030124510A1
GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R

Yarbough, Patricia O

Bradley, Daniel W

Krawczynski, Krzysztof Z

Tam, Albert

Fry, Kirk E

TITLE OF INVENTION: DNA Sequences of Enterically Transmitted

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS: No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Virus

ADDRESSEE: Dellinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/851,410

FILING DATE: 07-May-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/128,275

FILING DATE: <unknown>

APPLICATION NUMBER: US 07/681,078

FILING DATE: 05-Apr-1991

APPLICATION NUMBER: US 07/505,888

FILING DATE: 05-Apr-1990

APPLICATION NUMBER: US 07/420,921

FILING DATE: 13-Oct-1989

APPLICATION NUMBER: US 07/367,486

FILING DATE: 16-Jun-1989

APPLICATION NUMBER: US 07/336,672

FILING DATE: 11-Apr-1989

APPLICATION NUMBER: US 07/208,997

FILING DATE: 17-Jun-1988

ATTORNEY/AGENT INFORMATION:

NAME: Pelthory, Joanne R.

REGISTRATION NUMBER: 42,995

REFERENCE/DOCKET NUMBER: 4600-0183.24

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1693 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-851-410-7

Query Match 90.0%; Score 27; DB 9; Length 1693;
Best Local Similarity 80.0%; Pred. No. 2.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
1:111
DB 431 YAOCR 435

RESULT 4

US-09-468-147-91
Sequence 91, Application US/09468147A
Publication No. US20030049601A1
GENERAL INFORMATION:

APPLICANT: Abbott Laboratories

Erker, James G.

APPLICANT: Erker, James G.

APPLICANT: Desai, Suresh M.

APPLICANT: Dawson, George J.

APPLICANT: Mushahwar, I. K.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING

FILE REFERENCE: 6232.US.P1

CURRENT APPLICATION NUMBER: US/09/468,147A

CURRENT FILING DATE: 1999-12-21

EARLIER APPLICATION NUMBER: US 09/173,141

EARLIER FILING DATE: 1998-10-15

EARLIER APPLICATION NUMBER: US 60/061,199

EARLIER FILING DATE: 1997-10-15

NUMBER OF SEQ ID NOS: 258

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 91

LENGTH: 1698

TYPE: PRT

ORGANISM: Hepatitis E Virus

FEATURE:

OTHER INFORMATION: xaa - Unknown or Other at position 174

OTHER INFORMATION: xaa - Unknown or Other at position 363

OTHER INFORMATION: xaa - Unknown or Other at position 1088

OTHER INFORMATION: xaa - Unknown or Other at position 1131

OTHER INFORMATION: xaa - Unknown or Other at position 1217

OTHER INFORMATION: xaa - Unknown or Other at position 1389

US-09-468-147-91

Query Match 90.0%; Score 27; DB 9; Length 1698;

Best Local Similarity 80.0%; Pred. No. 2.6e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
1:111
DB 422 YAOCR 426

US-09-468-147-166

Sequence 166, Application US/09468147A

Publication No. US20030049601A1

GENERAL INFORMATION:

APPLICANT: Abbott Laboratories

APPLICANT: Schlaunder, George G.

APPLICANT: Erker, James G.

APPLICANT: Desai, Suresh M.

```

: APPLICANT: Dawson, George J.
: APPLICANT: Mushahwar, I. K.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
: TITLE OF INVENTION: HEPATITIS E VIRUS
: FILE REFERENCE: 6232.US.P1
: CURRENT APPLICATION NUMBER: US/09/468,147A
: CURRENT FILING DATE: 1999-12-21
: EARLIER APPLICATION NUMBER: US 09/173,141
: EARLIER FILING DATE: 1998-10-15
: EARLIER APPLICATION NUMBER: US 60/061,199
: EARLIER FILING DATE: 1997-10-15
: NUMBER OF SEQ ID NOS: 258
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 166
: LENGTH: 1708
: TYPE: PRT
: ORGANISM: Hepatitis E Virus
: FEATURE:
: OTHER INFORMATION: Xaa - Unknown or Other at position 322
: OTHER INFORMATION: Xaa - Unknown or Other at position 331
: OTHER INFORMATION: Xaa - Unknown or Other at position 445
: OTHER INFORMATION: Xaa - Unknown or Other at position 448
: OTHER INFORMATION: Xaa - Unknown or Other at position 634
: OTHER INFORMATION: Xaa - Unknown or Other at position 646
: OTHER INFORMATION: Xaa - Unknown or Other at position 811
: OTHER INFORMATION: Xaa - Unknown or Other at position 1553
: OTHER INFORMATION: Xaa - Unknown or Other at position 1578
: OTHER INFORMATION: Xaa - Unknown or Other at position 1691
: US-09-468-147-166

```

```

Query Match          90.0%; Score 27; DB 9; Length 1708;
Best Local Similarity 80.0%; Pred. No. 2.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 YSCR 5
    1:1111
Db 431 YACR 435

```

```

RESULT 6
: US-09-864-761-42491
: Sequence 42491, Application US/09864761
: Patent No. US2002048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30

```

```

: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 42491
: LENGTH: 103
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL139289.4
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
: OTHER INFORMATION: EST HUMAN HIT: AW87776.1, EVALUATE 2.00e+43
: OTHER INFORMATION: SWISSPROT HIT: O0839, EVALUATE 1.80e+00
: US-09-864-761-42491

```

```

Query Match          86.7%; Score 26; DB 10; Length 103;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 YSCR 5
    1:1111
Db 76 YSCR 80

```

```

RESULT 7
: US-09-925-300-1551
: Sequence 1551, Application US/09925300
: Patent No. US20020151681A1
: GENERAL INFORMATION:
: APPLICANT: Craig Rosen,
: APPLICANT: Steve Ruben,
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA101
: CURRENT APPLICATION NUMBER: US/09/925,300
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05988
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1890
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 1551
: LENGTH: 114
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-925-300-1551

```

```

Query Match          86.7%; Score 26; DB 10; Length 114;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 YSCR 5

```

Db 103 FSOCR 107

RESULT 8
US-09-734-569-130

; Sequence 130, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Benz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 130
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-130

Query Match 86.7%; Score 26; DB 10; Length 199;

Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 56 YSOCR 60

RESULT 9

US-09-841-132-20
; Sequence 20, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhalla, Ajay
; APPLICANT: Skelly, Yael A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 20
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-841-132-20

Query Match 86.7%; Score 26; DB 10; Length 216;

Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 102 YSOCR 106

RESULT 10
US-09-925-300-1601

; Sequence 1601, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruden,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1601
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1601

Query Match 86.7%; Score 26; DB 10; Length 253;

Best Local Similarity 80.0%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 109 YSOCR 113

RESULT 11

US-09-738-626-5525
; Sequence 5525, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5525
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5525

Query Match 86.7%; Score 26; DB 9; Length 289;

Best Local Similarity 80.0%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 265 YSOCR 269

RESULT 12
US-09-962-276-2
Sequence 2, Application US/09962276
Patent No. US20020025570A1
GENERAL INFORMATION:
APPLICANT: YE, Jane et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CLO01156DIV
CURRENT APPLICATION NUMBER: US/09/962,276
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/210,458
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 09/799,345
PRIOR FILING DATE: 2001-03-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 295
TYPE: PRT
ORGANISM: Human
US-09-962-276-2

Query Match 86.7%; Score 26; DB 10; Length 295;
Best Local Similarity 80.0%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YSQCR 5
| | | |
Db 208 YQOCR 212

RESULT 13
US-10-007-521-10
Sequence 10, Application US/10007521
Publication No. US20030054539A1
GENERAL INFORMATION:
APPLICANT: Schulein, Martin
Andersen, Lene N.
Lassen, Soren F.
Kauppinen, Markus S.
Lange, Lene
Nielsen, Ruby I.
Ihara, Michiko
Takagi, Shinobu
TITLE OF INVENTION: NO. US20030054539A1 Endoglucanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. US20030054539A1 No. US20030054539A1disk of No. US20030054539A1
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366,200-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-007-521-10

Query Match 86.7%; Score 26; DB 9; Length 349;
Best Local Similarity 80.0%; Pred. No. 9.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQCR 5
| | | |
Db 339 YSQCR 343

RESULT 14
US-10-156-761-8811
Sequence 8811, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATORI, MASAHIRO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 8811
LENGTH: 356
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-8811

Query Match 86.7%; Score 26; DB 9; Length 356;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YSQCR 5
| | | |
Db 101 YQOCR 105

RESULT 15
US-10-104-580-8
Sequence 8, Application US/10104580
Publication No. US2003003628A1
GENERAL INFORMATION:
APPLICANT: Strauss et al.
TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
poplar and other plant species.
FILE REFERENCE: 62486
CURRENT APPLICATION NUMBER: US/10/104,580
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 09/410,464
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 09/287,700
PRIOR FILING DATE: 1999-04-06
PRIOR APPLICATION NUMBER: 60/080,851

; PRIOR FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 8
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Populus balsamifera subsp. trichocarpa
US-10-104-580-8

Query Match 86.7%; Score 26; DB 9; Length 377;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YSQR 5
1 111
Db 233 YEQCR 237

Search completed: July 9, 2003, 12:17:58
Job time : 20.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:00:11 ; Search time 143.25 Seconds
(without alignments)
22,504 Million cell updates/sec

Title: US-09-851-058-2
Perfect score: 30
Sequence: 1 YSQR 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US111_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	35	US-60-136-244-154	Sequence 154, App
2	30	100.0	54	US-60-142-896-1326	Sequence 1326, Ap
3	30	100.0	54	US-60-143-992-1191	Sequence 1191, Ap
4	30	100.0	61	US-09-617-681A-8383	Sequence 8383, Ap
5	30	100.0	87	US-60-311-261-2772	Sequence 2772, Ap
6	30	100.0	119	US-09-617-681A-8381	Sequence 8381, Ap

7	30	100.0	132	US-60-140-956-2061	Sequence 2061, Ap
8	30	100.0	208	US-60-142-896-1526	Sequence 1526, Ap
9	30	100.0	214	PCT-US01-14827-10486	Sequence 10486, A
10	30	100.0	243	US-10-027-450-22	Sequence 22, Appl
11	30	100.0	280	US-10-219-999-56114	Sequence 56114, A
12	30	100.0	368	US-09-708-427-17468	Sequence 17468, A
13	30	100.0	371	US-09-252-991A-19630	Sequence 19630, A
14	30	100.0	377	US-09-708-427-17467	Sequence 17467, A
15	30	100.0	516	US-09-708-427-17466	Sequence 17466, A
16	30	100.0	729	US-60-360-039-6201	Sequence 6201, Ap
17	27	90.0	97	US-09-513-996A-11153	Sequence 11153, A
18	27	90.0	52	PCT-US00-2652AB-7533	Sequence 7533, Ap
19	27	90.0	52	US-10-106-698-7543	Sequence 7543, Ap
20	27	90.0	82	US-09-270-767-62128	Sequence 62128, A
21	27	90.0	82	US-09-270-849B-194520	Sequence 194520, A
22	27	90.0	84	US-09-417-507-28823	Sequence 28823, A
23	27	90.0	92	PCT-US01-08631-47225	Sequence 47225, A
24	27	90.0	125	US-09-417-507-34621	Sequence 34621, A
25	27	90.0	126	US-09-417-507-37015	Sequence 37015, A
26	27	90.0	149	US-09-540-209B-6718	Sequence 6718, Ap
27	27	90.0	151	US-09-675-784A-10022	Sequence 10022, A
28	27	90.0	153	PCT-US01-08631-39659	Sequence 39659, A
29	27	90.0	155	US-09-791-537-101157	Sequence 101157, A
30	27	90.0	155	US-09-791-537-12918	Sequence 12918, A
31	27	90.0	155	US-09-791-537-129201	Sequence 129201, A
32	27	90.0	155	US-09-791-537-129589	Sequence 129589, A
33	27	90.0	157	US-09-791-537-153532	Sequence 153532, A
34	27	90.0	162	US-10-029-386-29676	Sequence 29676, A
35	27	90.0	164	US-09-902-540-14545	Sequence 14545, A
36	27	90.0	187	PCT-US01-08631-43487	Sequence 43487, A
37	27	90.0	202	US-09-543-681A-7996	Sequence 7996, Ap
38	27	90.0	221	US-09-902-540-15515	Sequence 15515, A
39	27	90.0	252	US-09-270-767-46535	Sequence 46535, A
40	27	90.0	282	US-60-360-039-14452	Sequence 14452, A
41	27	90.0	311	US-09-540-209B-3549	Sequence 3549, Ap
42	27	90.0	330	US-09-614-150-2493	Sequence 2493, Ap
43	27	90.0	320	US-60-167-211-2541	Sequence 2541, Ap
44	27	90.0	330	US-60-173-466-2058	Sequence 2058, Ap
45	27	90.0	320	US-60-191-637-2499	Sequence 2499, Ap

ALIGNMENTS

RESULT 1
US-60-136-244-154
Sequence 134, Application US/60136244
GENERAL INFORMATION:
APPLICANT: Kerlavage, Anthony
TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
FILE REFERENCE: C1000019
CURRENT APPLICATION NUMBER: US/60/136,244
CURRENT FILING DATE: 1999-05-26
NUMBER OF SEQ ID NOS: 176
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 154
LENGTH: 35
TYPE: PRT
ORGANISM: Drosophila
US-60-136-244-154

Query Match 100.0%; Score 30; DB 27; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
Db 14 YSQR 18
RESULT 2
US-60-142-896-1326

Sequence 1326, Application US/60142896
GENERAL INFORMATION:
APPLICANT: Kerlavage, Anthony
TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
FILE REFERENCE: C1000052
CURRENT FILING DATE: 1999-07-08
CURRENT APPLICATION NUMBER: US/60/142,896
NUMBER OF SEQ ID NOS: 1584
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1326
LENGTH: 54
TYPE: PRT
ORGANISM: Drosophila
US-60-142-896-1326

Query Match
Best Local Similarity 100.0%; Score 30; DB 27; Length 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
DB 11 YSOGR 15

RESULT 3
US-60-143-992-1191
Sequence 1191, Application US/60143992
GENERAL INFORMATION:
APPLICANT: Kerlavage, Anthony
TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
FILE REFERENCE: C1000057
CURRENT FILING DATE: 1999-07-15
CURRENT APPLICATION NUMBER: US/60/143,992
NUMBER OF SEQ ID NOS: 1536
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1191
LENGTH: 54
TYPE: PRT
ORGANISM: Drosophila
US-60-143-992-1191

Query Match
Best Local Similarity 100.0%; Score 30; DB 27; Length 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
DB 11 YSOGR 15

RESULT 4
US-09-617-681A-8383
Sequence 8383, Application US/09617681A
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1064P
CURRENT FILING DATE: 2000-07-19
CURRENT APPLICATION NUMBER: US/09/617,681A
NUMBER OF SEQ ID NOS: 9311
SEQ ID NO 8383
LENGTH: 61
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..61
OTHER INFORMATION: Ceres Seq. ID 1439795
LOCATION: 1..61

OTHER INFORMATION: Xaa is any amino acid
US-09-617-681A-8383

Query Match
Best Local Similarity 100.0%; Score 30; DB 20; Length 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
DB 48 YSOGR 52

RESULT 5
US-60-311-261-2772
Sequence 2772, Application US/60311261
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Zhilwei
APPLICANT: Meng, Gezhi
APPLICANT: Ma, Yungqing
TITLE OF INVENTION: Novel Nucleic Acids and Secreted Polypeptides
FILE REFERENCE: 806
CURRENT APPLICATION NUMBER: US/60/311,261
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/577,409
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: US 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 09/552,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/770,160
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: US 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 2818
SOFTWARE: PL-FL-genes Version 5.0
SEQ ID NO 2772
LENGTH: 87
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(87)
OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the Table 8 as set

US-60-311-261-2772

Query Match
Best Local Similarity 100.0%; Score 30; DB 27; Length 87;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
DB 58 YSOGR 62

```
RESULT 6
US-09-617-681A-8381
; Sequence 8381, Application US/09617681A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1064P
; CURRENT APPLICATION NUMBER: US/09/617,681A
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 9311
; SEQ ID NO 8381
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..119
; OTHER INFORMATION: Ceres Seq. ID 1439793
; NAME/KEY: misc-feature
; LOCATION: 1..119
; OTHER INFORMATION: Xaa is any amino acid
US-09-617-681A-8381

Query Match
Best Local Similarity 100.0%; Score 30; DB 20; Length 119;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
    |||||
Db 106 YSOCR 110

RESULT 7
US-60-140-956-2061
; Sequence 2061, Application US/60140956
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: C1000041
; CURRENT APPLICATION NUMBER: US/60/140,956
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 2638
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2061
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Drosophila
US-60-140-956-2061

Query Match
Best Local Similarity 100.0%; Score 30; DB 27; Length 132;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
    |||||
Db 15 YSOCR 19

RESULT 8
US-60-142-896-1526
; Sequence 1526, Application US/60142896
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: C1000052
; CURRENT APPLICATION NUMBER: US/60/142,896
; CURRENT FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 1584
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 1526
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Drosophila
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(208)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-142-896-1526

Query Match
Best Local Similarity 100.0%; Score 30; DB 27; Length 208;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
    |||||
Db 109 YSOCR 113

RESULT 9
PCT-US01-14827-10486
; Sequence 10486, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 10486
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (43)...(64)
; OTHER INFORMATION: GPI/DBG GTP-BINDING PROTEIN FAMILY SIGNATURE domain
; OTHER INFORMATION: Identified by EMATRIX, accession number PR00336A, p-value-6.25
PCT-US01-14827-10486

Query Match
Best Local Similarity 100.0%; Score 30; DB 1; Length 214;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
    |||||
Db 6 YSOCR 10

RESULT 10
US-10-027-450-22
; Sequence 22, Application US/10027450
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/10/027,450
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/063,423
; PRIOR FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 22
; LENGTH: 243
; TYPE: PRT
```

ORGANISM: Glycine max
US-10-027-450-22

Query Match 100.0%; Score 30; DB 24; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
DB 68 YSQR 72

RESULT 11

US-10-219-999-56114
; Sequence 56114, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kowalik, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 56114
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Zea mays
US-10-219-999-56114

Query Match 100.0%; Score 30; DB 26; Length 280;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
DB 32 YSQR 36

RESULT 12

US-09-708-427-17468
; Sequence 17468, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17468
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..368
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..368
; OTHER INFORMATION: Ceres Seq. ID 1832395
US-09-708-427-17468

Query Match 100.0%; Score 30; DB 21; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
DB 182 YSQR 186

RESULT 13

US-09-252-991A-19630
; Sequence 19630, Application US/09252991A
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19630
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19630

Query Match 100.0%; Score 30; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
DB 115 YSQR 119

RESULT 14

US-09-708-427-17467
; Sequence 17467, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17467
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..377
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..377
; OTHER INFORMATION: Ceres Seq. ID 1832394
US-09-708-427-17467

Query Match 100.0%; Score 30; DB 21; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
DB 191 YSQR 195

RESULT 15
US-09-708-427-17466

```
; Sequence 17466, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17466
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..516
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..516
; OTHER INFORMATION: Ceres Seq. ID 1832393
; US-09-708-427-17466
```

```
Query Match          100.0%; Score 30; DB 21; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 YSOCR 5
        |||||
Db       330 YSOCR 334
```

```
Search completed: July 9, 2003, 12:13:17
Job time : 145.25 secs
```

This Page Blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:00:51 ; Search time 47.25 Seconds
(Without alignments)
27.740 Million cell updates/sec

Title: US-09-851-058-2

Perfect score: 30
Sequence: 1 YSCGR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1221054 seqs, 262139423 residues

Total number of hits satisfying chosen parameters: 1221054

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCr_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCr_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	5	2 PCT-US02-14369-2	Sequence 2, Appl1
2	30	100.0	5	10 US-09-851-058-2	Sequence 2, Appl1
3	30	100.0	60	12 US-10-424-599-145972	Sequence 145972,
4	30	100.0	87	12 US-10-219-382-2772	Sequence 2772, Ap
5	30	100.0	87	12 US-10-219-382A-2772	Sequence 2772, Ap
6	30	100.0	108	12 US-10-424-599-229646	Sequence 229646,
7	30	100.0	280	12 US-10-425-114-64966	Sequence 64966, A
8	30	100.0	332	12 US-10-437-963-173488	Sequence 173488,
9	30	100.0	371	12 US-10-366-683-19630	Sequence 19630, A
10	30	100.0	371	12 US-10-419-128-19630	Sequence 19630, A
11	30	100.0	729	12 US-10-369-493-6201	Sequence 6201, Ap
12	27	90.0	5	2 PCT-US02-14369-3	Sequence 3, Appl1
13	27	90.0	5	10 US-09-851-058-3	Sequence 3, Appl1
14	27	90.0	38	12 US-10-424-599-157300	Sequence 157300,
15	27	90.0	83	12 US-10-424-599-165725	Sequence 165725,
16	27	90.0	85	12 US-10-424-599-180856	Sequence 180856,
17	27	90.0	109	12 US-10-424-599-171831	Sequence 171831,
18	27	90.0	124	10 US-09-864-408A-292	Sequence 292, App
19	27	90.0	124	12 US-10-218-140-574	Sequence 574, App

20	27	90.0	151	10 US-09-675-784A-10022	Sequence 10022, A
21	27	90.0	155	12 US-10-219-051B-12561	Sequence 12561, A
22	27	90.0	155	12 US-10-219-051B-14320	Sequence 14320, A
23	27	90.0	292	12 US-10-369-493-12452	Sequence 12452, A
24	27	90.0	304	12 US-10-282-122A-57028	Sequence 57028, A
25	27	90.0	372	12 US-10-437-963-157136	Sequence 157136,
26	27	90.0	458	12 US-10-438-246-17076	Sequence 17076, A
27	27	90.0	539	12 US-10-438-246-17076	Sequence 17076, A
28	27	90.0	539	12 US-10-438-246-24507	Sequence 24507, A
29	27	90.0	555	12 US-10-369-493-12740	Sequence 12740, A
30	27	90.0	582	12 US-10-437-963-147204	Sequence 147204,
31	27	90.0	676	12 US-10-369-493-3268	Sequence 3268, Ap
32	27	90.0	692	12 US-10-231-222B-38	Sequence 38, Appl
33	27	90.0	1693	2 PCT-US02-36096-5	Sequence 5, Appl1
34	27	90.0	1698	12 US-10-319-745-91	Sequence 91, Appl
35	27	90.0	1708	12 US-10-319-745-166	Sequence 166, App
36	27	90.0	2275	14 US-60-452-680-23634	Sequence 23634, A
37	27	86.7	45	12 US-10-219-382-1089	Sequence 1089, Ap
38	26	86.7	63	12 US-10-219-382A-1089	Sequence 1089, Ap
39	26	86.7	63	12 US-10-424-599-200237	Sequence 200237,
40	26	86.7	65	12 US-10-437-963-132751	Sequence 132751,
41	26	86.7	66	12 US-10-437-963-103103	Sequence 103103,
42	26	86.7	67	12 US-10-424-599-235524	Sequence 235524,
43	26	86.7	69	2 PCT-US02-32727-6297	Sequence 6297, Ap
44	26	86.7	69	10 US-09-978-825-6297	Sequence 6297, Ap
45	26	86.7	69	12 US-10-057-498-6297	Sequence 6297, Ap

ALIGNMENTS

RESULT 1
PCT-US02-14369-2
Sequence 2, Application PC/TUS0214369
GENERAL INFORMATION:
APPLICANT: Parker, Kenneth
APPLICANT: Nadler, Timothy
APPLICANT: Vella, George
APPLICANT: Huang, Yulin
APPLICANT: Abersold, Rudolf
APPLICANT: Smolka, Marcus
TITLE OF INVENTION: Process for Analyzing Protein Samples
FILE REFERENCE: STP-172
CURRENT APPLICATION NUMBER: PCT/US02/14369
CURRENT FILING DATE: 2003-02-05
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent version 3.1
SEQ ID NO 2
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptides
PCT-US02-14369-2
Query Match 100.0%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1;le+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 YSCGR 5
Y 1 YSCGR 5
DB 1 YSCGR 5
RESULT 2
US-09-851-058-2
Sequence 2, Application US/09851058
GENERAL INFORMATION:
APPLICANT: Parker, Kenneth
APPLICANT: Nadler, Timothy
APPLICANT: Vella, George
APPLICANT: Huang, Yulin
APPLICANT: Abersold, Rudolf

APPLICANT: Smolke, Marcus
TITLE OF INVENTION: Process for Analyzing Protein Samples
FILE REFERENCE: SYP-172
CURRENT APPLICATION NUMBER: US/09/851,058
CURRENT FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent version 3.1
SEQ ID NO 2
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptides
US-09-851-058-2

Query Match 100.0%; Score 30; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.1e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
DB 1 YSQR 5

RESULT 3
US-10-424-599-145972
SEQUENCE 145972, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 145972
LENGTH: 60
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(60)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_102831C.1.pcp
US-10-424-599-145972

Query Match 100.0%; Score 30; DB 12; Length 60;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
DB 51 YSQR 55

RESULT 4
US-10-219-382-2772
SEQUENCE 2772, Application US/10219382
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Zhilwei
APPLICANT: Meng, Gezhi
APPLICANT: Ma, Yungqing
TITLE OF INVENTION: Novel Nucleic Acids and Secreted Polypeptides
FILE REFERENCE: 806A
CURRENT APPLICATION NUMBER: US/10/219,382
CURRENT FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,261

PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/577,409
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/574,454
Remaining prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2818
SOFTWARE: PL-FL-genes Version 5.0
SEQ ID NO 2772
LENGTH: 87
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(87)
OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the Table 8 as set
US-10-219-382-2772

Query Match 100.0%; Score 30; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
DB 58 YSQR 62

RESULT 5
US-10-219-382A-2772
SEQUENCE 2772, Application US/10219382A
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Zhilwei
APPLICANT: Meng, Gezhi
APPLICANT: Ma, Yungqing
TITLE OF INVENTION: Novel Nucleic Acids and Secreted Polypeptides
FILE REFERENCE: 806A
CURRENT APPLICATION NUMBER: US/10/219,382A
CURRENT FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,261
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/577,409
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/574,454

PRIOR FILING DATE: 2000-05-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2818
; SOFTWARE: PL_FL_genes Version 5.0
; SEQ ID NO 2772
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(87)
; OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the Table 8 as set for
US-10-219-382A-2772

Query Match 100.0%; Score 30; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
Db 58 YSQR 62

RESULT 6
US-10-424-599-229646
; Sequence 229646, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovall David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 229646
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(108)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49395C.1.pep
US-10-424-599-229646

Query Match 100.0%; Score 30; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
Db 100 YSQR 104

RESULT 7
US-10-425-114-64966
; Sequence 64966, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovall, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53137)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64966
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4732-009-A1_FL1.pep
US-10-425-114-64966

Query Match 100.0%; Score 30; DB 12; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
Db 32 YSQR 36

RESULT 8
US-10-437-963-173488
; Sequence 173488, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovall, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173488
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71520C.1.pep
US-10-437-963-173488

Query Match 100.0%; Score 30; DB 12; Length 332;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSQR 5
|||||
Db 157 YSQR 161

RESULT 9
US-10-366-683-19630
; Sequence 19630, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolling, Jork
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERIGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PAT03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19630
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-10-366-683-19630

Query Match 100.0%; Score 30; DB 12; Length 371;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
DB 115 YSOGR 119

RESULT 10
US-10-419-128-19630

; Sequence 19630, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.116
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19630
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-19630

Query Match 100.0%; Score 30; DB 12; Length 371;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
DB 115 YSOGR 119

RESULT 11
US-10-369-493-6201
; Sequence 6201, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6201
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6201

Query Match 100.0%; Score 30; DB 12; Length 729;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
DB 536 YSOGR 540

RESULT 12
PCT-US02-14369-3

; Sequence 3, Application PC/YUS0214369
; GENERAL INFORMATION:
; APPLICANT: Parker, Kenneth
; APPLICANT: Nadler, Timothy
; APPLICANT: Vella, George
; APPLICANT: Huang, Yulin
; APPLICANT: Adersold, Rudolf
; APPLICANT: Smolka, Marcus
; TITLE OF INVENTION: Process for Analyzing Protein Samples
; FILE REFERENCE: SYP-172
; CURRENT APPLICATION NUMBER: PCT/US02/14369
; CURRENT FILING DATE: 2003-02-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptides
PCT-US02-14369-3

Query Match 90.0%; Score 27; DB 2; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.1e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
DB 1 YSOGR 5

RESULT 13
US-09-851-058-3
; Sequence 3, Application US/09851058
; GENERAL INFORMATION:
; APPLICANT: Parker, Kenneth
; APPLICANT: Nadler, Timothy
; APPLICANT: Vella, George
; APPLICANT: Huang, Yulin
; APPLICANT: Adersold, Rudolf
; APPLICANT: Smolka, Marcus
; TITLE OF INVENTION: Process for Analyzing Protein Samples
; FILE REFERENCE: SYP-172
; CURRENT APPLICATION NUMBER: US/09/851,058
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptides
US-09-851-058-3

Query Match 90.0%; Score 27; DB 10; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.1e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
DB 1 YSOGR 5

RESULT 14
US-10-424-599-157300
; Sequence 157300, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 157300
LENGTH: 38
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_113062C.1.pap
US-10-424-599-157300

Query Match 90.0%; Score 27; DB 12; Length 38;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
1:111
Db 28 YAOGR 32

RESULT 15
US-10-424-599-165725
Sequence 165725, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 165725
LENGTH: 83
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_120665C.1.pap
US-10-424-599-165725

Query Match 90.0%; Score 27; DB 12; Length 83;
Best Local Similarity 80.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
1:111
Db 16 YSECR 20

Search completed: July 9, 2003, 12:16:32
Job time : 49.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:58:06 ; Search time 9.25 Seconds
(Without alignments)
51.965 Million cell updates/sec

Title: US-09-851-058-2

Perfect score: 30

Sequence: 1 YSOGR 5

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:**
1: PIR1:**
2: PIR2:**
3: PIR3:**
4: PIR4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	221	2 T20781	hypothetical prote
2	30	100.0	341	2 B82790	conserved hypothet
3	30	100.0	516	2 G84442	probable nucleosid
4	30	100.0	729	2 T23474	hypothetical prote
5	27	90.0	78	2 T17891	hypothetical prote
6	27	90.0	155	2 JC6159	eosinophil-associa
7	27	90.0	156	2 JC6160	transcription regu
8	27	90.0	315	2 F82135	chorismate mutase
9	27	90.0	357	2 H81450	hypothetical prote
10	27	90.0	378	2 B85059	hypothetical prote
11	27	90.0	451	2 A70539	hypothetical prote
12	27	90.0	501	2 A70534	hypothetical prote
13	27	90.0	692	2 T13161	A-kinase anchor pr
14	27	90.0	860	2 T13778	hypothetical prote
15	27	90.0	1171	2 T12956	hypothetical prote
16	27	90.0	1691	1 A44212	genome polypotein
17	27	90.0	1693	1 MNMME	hypothetical prote
18	26	86.7	121	2 T16049	hypothetical prote
19	26	86.7	127	2 A86155	hypothetical prote
20	26	86.7	185	1 R52M64	ribosomal protein
21	26	86.7	185	2 T03233	probable CDP-alcoh
22	26	86.7	188	2 E71157	hypothetical prote
23	26	86.7	201	2 AG2434	probable DNA-bind
24	26	86.7	212	2 G98143	splicing factor SR
25	26	86.7	221	2 S59075	probable membrane
26	26	86.7	259	2 G97188	hypothetical prote
27	26	86.7	298	2 S27965	probable transcript
28	26	86.7	298	2 D83319	2-phosphoglycerate
29	26	86.7	304	2 S41582	

30	26	86.7	306	1 A69119	2-phosphoglycerate
31	26	86.7	309	2 AH0946	probable lipase ST
32	26	86.7	324	2 H87540	alcohol dehydrogen
33	26	86.7	326	2 G64894	hypothetical prote
34	26	86.7	346	2 F96695	hypothetical prote
35	26	86.7	372	2 H97277	glycosyltransferas
36	26	86.7	388	1 JC5461	cellulase (EC 3.2.
37	26	86.7	388	1 S43920	cellulase (EC 3.2.
38	26	86.7	389	2 T03411	protein RFL - rice
39	26	86.7	395	2 T06247	UNIPOLIANA protein
40	26	86.7	395	2 T06372	probable transcrip
41	26	86.7	396	2 A16339	FLD protein - gard
42	26	86.7	412	2 S76239	hypothetical prote
43	26	86.7	412	2 C38104	LFY floral meriste
44	26	86.7	413	2 T03240	FLD/LFY protein ho
45	26	86.7	415	2 S97340	flto protein homolo

ALIGNMENTS

RESULT 1

T20781 hypothetical protein F11D11.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T20781

R:Mortimore, B.; Basham, V.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19323

A:Accession: T20781

A>Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-221 <MIL>

A:Cross-references: EMBL:Z81500; PIDN:CA804099.1; GSPDB:GN00023; CESP:F11D11.8

A:Experimental source: clone F11D11

C:Genetics:

A:Gene: CESP:F11D11.8

A:Map position: 5

A:Insertions: 69/3

Query Match Best Local Similarity 100.0%; Score 30; DB 2; Length 221;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5

Db 107 YSOGR 111

RESULT 2

B82790

conserved hypothetical protein XF0556 (Imported) - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82790

R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: B82790

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-341 <STW>

A:Cross-references: GB:AE003903; GB:AE003849; NID:9105416; PIDN:AF83366.1; GSPDB:GN

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to Genbank, June 2000

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klieger, J.E.; Kuramae, E.E.; La

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, H.A.; Nogueira, Martins, E.M.F.; Matsushima, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0556

Query Match 100.0%; Score 30; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 34 YSOCR 38

RESULT 3

probable nucleoside triphosphatase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84442

R:Lin, X.; Kuhl, S.; Rounsley, S.D.; Shea, T.P.; Beattie, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euns, D.; Nierman, W.C.; White, O.; Eskin, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MID:20083487; PMID:10617197
A:Accession: G84442
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-516 <STO>
A:Cross-references: GB:AEO02093; NID:g3461821; PIDN:AAC32915.1; GSPDB:GN00139
C:Genetics:
A:Gene: Atg292970
A:Map position: 2

Query Match 100.0%; Score 30; DB 2; Length 516;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 330 YSOCR 334

RESULT 4

hypothetical protein K08E7.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23474
R:Smyle, R.

submitted to the EMBL Data Library, July 1996
A:Reference number: Z19745
A:Accession: T23474
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-729 <WIL>
A:Cross-references: EMBL:Z77666; PIDN:CAH01230.1; GSPDB:GN00022; CESP:K08E7.7
A:Experimental source: clone K08E7
C:Genetics:
A:Gene: CESP:K08E7.7
A:Map position: 4
A:introns: 33/2; 292/2; 405/3; 555/3; 665/3; 686/3

Query Match 100.0%; Score 30; DB 2; Length 729;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 536 YSOCR 540

RESULT 5

hypothetical protein a388r - Chlorella virus PCV-1
C:Species: Chlorella virus PCV-1
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17891
R:Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T17891
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-78 <GRA>
A:Cross-references: EMBL:U42580; NID:9402896; PIDN:AAC96756.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: a388r

Query Match 90.0%; Score 27; DB 2; Length 78;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 27 YSOCR 31

RESULT 6

eosinophil-associated ribonuclease (EC 3.1.-.-) 1 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 22-Jun-1999
C:Accession: JG6159; PC6041
R:Larson, K.A.; Olson, E.V.; Madden, B.J.; Gleich, G.J.; Lee, N.A.; Lee, J.J.

Proc. Natl. Acad. Sci. U.S.A. 93, 12370-12375, 1996
A:Title: Two highly homologous ribonuclease genes expressed in mouse eosinophils
A:Reference number: JG6159; MID:97057249; PMID:8901588
A:Accession: JG6159
A:Molecule type: mRNA
A:Residues: 1-155 <LAR1>
A:Cross-references: GB:U72032; NID:g1695898; PIDN:AAB37786.1; PID:g1695899
A:Accession: PC6041
A:Molecule type: protein
A:Residues: 52-75; 137-149 <LAR2>
A:Experimental source: bone marrow
C:Genetics:
A:Gene: mear-1
C:Superfamily: pancreatic ribonuclease
C:Keywords: bone; disulfide bond; glycoprotein; hydrolase; RNA binding
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-155/Product: eosinophil-associated ribonuclease 1 #status predicted <MAT>
F:38,62,150/Active site: His, Lys, His #status predicted
F:41,96,107,114/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:47-106,61-118,79-133,86-94/Disulfide bonds: #status predicted

Query Match 90.0%; Score 27; DB 2; Length 155;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
DB 115 YSOCR 119

RESULT 7

JG6160
eosinophil-associated ribonuclease (EC 3.1.-.-) 2 precursor - mouse

C:Species: Mus musculus (house mouse)
C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 22-Jun-1999
C:Accession: J06160; PC6042
R:Larson, K.A.; Olson, E.V.; Madden, B.J.; Gleish, G.J.; Lee, N.A.; Lee, J.J.
Proc. Natl. Acad. Sci. U.S.A. 93, 12370-12375, 1996
A:Title: Two highly homologous ribonuclease genes expressed in mouse eosinophils identified
A:Reference number: J06159; MUID:97057249; PMID:8901588
A:Accession: J06160
A:Molecule type: mRNA
A:Residues: 1-156 <LAR1>
A:Cross-references: GB:U72031; NID:g1695896; PIDN:AAB37785.1; PID:g1695897
A:Experimental source: Bone marrow
A:Accession: PC6042
A:Molecule type: Protein
A:Residues: 52-77,138-150 <LAR2>
C:Genetics:
A:Gene: mear-2
C:Superfamily: pancreatic ribonuclease
C:Keywords: bone; disulfide bond; glycoprotein; hydrolase; RNA binding
F:1-37/Domain: signal sequence #status predicted <SIG>
F:28-156/Product: eosinophil-associated ribonuclease 2 #status predicted <MAT>
F:38,62,151/Active site: His, Lys, His #status predicted
F:47-106,61-119,79-134,86-94/Disulfide bonds: #status predicted
F:89,96,107/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.0%; Score 27; DB 2; Length 156;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YSQR 5
1:1111
Db 116 YTCR 120

RESULT 8
F82135
transcription regulator LysR family VC1955 [Imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: F82135
R:Heideberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Charidson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: F82135
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-315 <HE1>
A:Cross-references: GB:A0004271; GB:A0003852; NID:g9656490; PIDN:AA95103.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1955
A:Map position: 1.
C:Superfamily: conserved hypothetical protein H11364

Query Match 90.0%; Score 27; DB 2; Length 315;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YSQR 5
1:1111
Db 68 YACR 72

RESULT 9
H81450
chorismate mutase (EC 5.4.99.5) / prephenate dehydratase (EC 4.2.1.51) Cj0316 [Imported]
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: H81450
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Bar
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: H81450
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-357 <PAR>
A:Cross-references: GB:A1139074; GB:A1111168; NID:g6967505; PIDN:CAB72783.1; PID:g696
A:Experimental source: serotype O2; strain NCTC 11168
C:Genetics:
A:Gene: pheA; Cj0316
C:Superfamily: phea bifunctional enzyme; prephenate dehydratase homology
C:Keywords: carbon-oxygen lyase; hydro-lyase; intramolecular transferase; isomerase

Query Match 90.0%; Score 27; DB 2; Length 357;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YSQR 5
1:1111
Db 193 YTCR 197

RESULT 10
B85059
hypothetical protein AT4g04690 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: B85059
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: B85059
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-378 <STO>
A:Cross-references: GB:NC_001268; NID:g7267227; PIDN:CAB80834.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g04690
A:Map position: 4

Query Match 90.0%; Score 27; DB 2; Length 378;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YSQR 5
1:1111
Db 194 YTCR 198

RESULT 11
A70539
hypothetical protein RV1128c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70539
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:96295987; PMID:9634230
A:Accession: A70539
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-451 <COL>
A:Cross-references: GB:g29585; GB:A1123456; NID:g3261787; PIDN:CAB09039.1; PID:g21172
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV1128c

C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV1945

Query Match 90.0%; Score 27; DB 2; Length 451;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCKR 5
DB 383 YSOCKR 387

RESULT 12

A70534

hypothetical protein RV0654 - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: A70534

R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon, S.

R:Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

R:Rejnders, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

R:Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: A70534

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-501 <COL>

A:Cross-references: GB:295972; GB:AL123456; NID:93261790; PIDN:CAB09380.1; PID:ea19166;

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: RV0654

Query Match 90.0%; Score 27; DB 2; Length 501;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCKR 5
DB 319 YSOCKR 323

RESULT 13

T13161

A:kinase anchor protein 95 - human

N:Alternate names: protein DKF2p586b1222.1

C:Species: Homo sapiens (man)

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C:Accession: T13161; T08786

R:Lamerding, J.E.; McCreedy, P.M.; Skowronski, E.; Viswanathan, V.; Burkhart-Schultz, K.

R:Quann, G.; Krommiller, B.; Erler, A.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Atlix

R:submitted to the EMBL Data Library, October 1998

A:Authors: Kobayashi, A.; Olsen, A.S.; Carrano, A.V.

A:Description: Sequence analysis of a 1.5 Mb OLF-R-rich region in 19p13.1.

A:Reference number: Z17617

A:Accession: T13161

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-692 <LAN>

A:Cross-references: EMBL:AC005785; NID:93702290; PID:93702291; PIDN:AAC62838.1

R:Roehner, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

R:submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16468

A:Accession: T08786

A:Molecule type: mRNA

A:Residues: 335-692 <KOE>

A:Cross-references: EMBL:AL050160

A:Experimental source: adult uterus; clone DKF2p586b1222

C:Genetics:

A:Map position: 19

A:Introns: 7/1; 20/1; 31/1; 124/2; 287/3; 331/1; 346/3; 358/1; 387/2; 434/3; 466/1; 509/

A:Note: DKF2p586b1222.1

Query Match 90.0%; Score 27; DB 2; Length 692;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCKR 5
DB 170 YSOCKR 174

RESULT 14

T15778

hypothetical protein C35B8.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000

C:Accession: T15778

R:Leimbach, D.

R:submitted to the EMBL Data Library, March 1995

A:Description: The sequence of C. elegans cosmid C35B8.

A:Reference number: Z18401

A:Accession: T15778

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-860 <LEI>

A:Cross-references: EMBL:U23520; NID:9746532; PID:9746534; PIDN:AAC46554.1; CESP:C35B

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C35B8.2

A:Introns: 37/2; 67/3; 126/2; 173/2; 200/3; 304/2; 372/3; 565/2; 615/1; 677/3; 714/2;

C:Superfamily: vav transforming protein; CDC24 homology; protein kinase C zinc-blndin

Query Match 90.0%; Score 27; DB 2; Length 860;
Best Local Similarity 80.0%; Pred. No. 5.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCKR 5
DB 853 YSOCKR 857

RESULT 15

T12956

hypothetical protein T6H20.50 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 22-Oct-1999

C:Accession: T12956

R:Choisne, N.; Robert, C.; Bottler, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.

R:submitted to the Protein Sequence Database, July 1999

A:Reference number: Z17586

A:Accession: T12956

A:Molecule type: DNA

A:Residues: 1-1171 <CHO>

A:Cross-references: EMBL:AL096859; GSPDB:GM00061; ATSP-T6H20.50

A:Experimental source: cultivar Columbia; BAC clone T6H20

C:Genetics:

A:Gene: ATSP-T6H20.50

A:Map position: 3

A:Introns: 780/1; 857/3; 932/3; 987/2; 1036/3; 1077/3; 1107/1

Query Match 90.0%; Score 27; DB 2; Length 1171;
Best Local Similarity 80.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCKR 5
DB 580 YSOCKR 584

Search completed: July 9, 2003, 12:02:45
Job time: 11.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:55:50 ; Search time 7.75 Seconds

(without alignments)
26.759 Million cell updates/sec

Title: US-09-851-058-2

Perfect score: 30
Sequence: 1 YSQR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	30	100.0	210	1	PSBP_PORPU
2	30	100.0	729	1	CUL6_CAEEL
3	27	90.0	155	1	ECPI_MOUSE
4	27	90.0	155	1	ECPI_MOUSE
5	27	90.0	156	1	ECPI_MOUSE
6	27	90.0	156	1	ECPI_MOUSE
7	27	90.0	451	1	YB28_MYCTU
8	27	90.0	692	1	AKA8_HUMAN
9	27	90.0	1691	1	POLN_HEYME
10	27	90.0	1693	1	POLN_HEYME
11	27	90.0	1693	1	POLN_HEYME
12	27	90.0	1693	1	POLN_HEYME
13	26	86.7	185	1	RM16_MAIZE
14	26	86.7	185	1	RM16_ORYSA
15	26	86.7	188	1	Y460_PYRHO
16	26	86.7	221	1	SER9_HUMAN
17	26	86.7	298	1	Y33K_HUMAN
18	26	86.7	326	1	YDCK_ECOLI
19	26	86.7	339	1	FLI_EUGL
20	26	86.7	377	1	FLH_POPTR
21	26	86.7	388	1	GUN3_HUMAN
22	26	86.7	389	1	FL_ORYSA
23	26	86.7	395	1	UNI_PEA
24	26	86.7	396	1	FLU_ANTMA
25	26	86.7	412	1	ALF_PETIR
26	26	86.7	413	1	FLI_TOBAC
27	26	86.7	415	1	BOFH_BRAOL
28	26	86.7	416	1	FL2_TOBAC
29	26	86.7	420	1	CREL_CHICK
30	26	86.7	424	1	LFT_ARATH
31	26	86.7	462	1	GUNB_FUSOX
32	26	86.7	514	1	GUXC_FUSOX
33	26	86.7	558	1	ORC2_XENLA

34	26	86.7	563	1	ENV_BAEYM	P10269 baboon endo
35	26	86.7	575	1	WGLE_HSVS	O08101 mouse herp
36	26	86.7	576	1	ORC2_MOUSE	O60602 mus musculu
37	26	86.7	577	1	ORC2_HUMAN	O13416 homo saplen
38	26	86.7	703	1	GCR_PIG	O9n133 sus scrofa
39	26	86.7	704	1	TRFL_PIG	P14632 sus scrofa
40	26	86.7	771	1	GCR_CAVPO	P49115 cavia porce
41	26	86.7	2248	1	CYAL_DROME	P32870 drosophila
42	26	86.7	2479	1	POLN_RRYN	P13887 ross river
43	26	86.7	4273	1	PKSL_BACSU	P40872 bacillus su
44	26	86.7	4427	1	PKSL_BACSU	O05470 bacillus su
45	25	83.3	44	1	YIM8_BPPH1	P10431 bacterioph

ALIGNMENTS

RESULT 1	PSBP_PORPU	STANDARD	PRT	210 AA.
ID	PSBP_PORPU			
AC	P50272			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	putative polysaccharide binding protein precursor.			
OS	Porphyra purpurea			
OC	Eukaryota; Rhodophyta; Bangiophyta; Bangiales; Bangiaceae; Porphyra			
OX	NCBI_Taxid-2787			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-Avonport			
RA	Lin O., der Meer J.P., Reith M.E.			
RL	Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases			
CC	-1 SIMILARITY: CONTAINS 4 FUNGAL-TYPE CELLULOSE-BINDING DOMAINS (CBD).			
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL: U08843; AAA61792.1; -			
DR	HSP; P00725; ZCBH.			
DR	InterPro: IPR00254; CBD_Fungal.			
DR	Pfam: PF00734; CBM_1; 4.			
DR	SMART: SM00236; FCBP; 4.			
DR	PROSITE: PS00562; CBD_FUNAL; 4.			
KW	Repeat: Signal.			
FT	SIGNAL	1	22	
FT	CHAIN	23	210	
FT	DOMAIN	23	62	PUTATIVE POLYSACCHARIDE BINDING PROTEIN.
FT	DOMAIN	63	105	POLYSACCHARIDE BINDING DOMAIN I.
FT	DOMAIN	125	165	POLYSACCHARIDE BINDING DOMAIN II.
FT	DOMAIN	166	210	POLYSACCHARIDE BINDING DOMAIN III.
FT	DOMAIN	166	210	POLYSACCHARIDE BINDING DOMAIN IV.
SO	SEQUENCE	210 AA; 21967 MW; 837820986863C8 CRC64;		
Query Match		Score 30; DB 1; Length 210;		
Best Local Similarity		100.0%; Pred. No. 18;		
Matches	5; Conservative	0; Mismatches	0; Indels	0; Caps
OY	1 YSQR 5			
DB	54 YSQR 58			
RESULT 2				
ID	CUL6_CAEEL	STANDARD	PRT	729 AA.
AC	Q21346			
DT	01-NOV-1997 (Rel. 35, Created)			

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CUL-6 protein.
 GN CUL-6 OR K08E7.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Smye R.;
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CULLIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z77666; CAB01230.1; -
 DR WormPep: K08E7.7; CE11928.
 DR InterPro: IPR001373; Cullin.
 DR Pfam: PF00888; Cullin. 1.
 DR SMART: SM00182; CULLIN. 1.
 DR PROSITE: PS01256; CULLIN_1; 1.
 DR PROSITE: PS50069; CULLIN_2; 1.
 DR SEQUENCE 729 AA; 84685 MW; 3D400932DE78E2CF CRC64;
 SO
 Query Match Best Local Similarity 100.0%; Score 30; DB 1; Length 729;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YSQCR 5
 DB 536 YSQCR 540
 RESULT 3
 ECP_MOUSE STANDARD; PRT; 155 AA.
 ID ECP_MOUSE
 AC P97426;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Eosinophil cationic protein 1 precursor (EC 3.1.27.-) (ECP 1)
 DE (Ribonuclease 3-1) (Rnase 3-1) (Eosinophil secondary granule
 DE ribonuclease-1) (EAR-1).
 GN EAR1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Larson K.A., Olson E.V., Madden B.J., Glatch G.J., Lee N.A., Lee J.J.;
 RT Two highly homologous ribonuclease genes expressed in mouse
 RT eosinophils identify a larger subgroup of the mammalian ribonuclease
 RT superfamily.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12370-12375(1996).
 CC -1- FUNCTION: CYTOTOXIN AND HELMINTHOTOXIN WITH RIBONUCLEASE ACTIVITY.
 CC POSSESSES A WIDE VARIETY OF BIOLOGICAL ACTIVITIES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MATRIX OF EOSINOPHIL'S LARGE SPECIFIC
 CC GRANULE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U72032; AB37786.1; -
 DR HSP: P00656; 2RNS.
 DR MGD: MG1:108021; Ear1.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; rnasea. 1.
 DR PRINTS: PR00794; RIBONUCLEASE.
 DR PRODOM: PD000535; RNaseA. 1.
 DR SMART: SM00092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNase_PANCREATIC; 1.
 DR Hydrolase: Nuclease; Endonuclease; Eosinophil; Toxin; Glycoprotein;
 KW Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 155
 FT ACT_SITE 38 38
 FT ACT_SITE 62 62
 FT ACT_SITE 150 150
 FT DISULFID 47 106
 FT DISULFID 61 118
 FT DISULFID 79 133
 FT DISULFID 86 94
 FT CARBOHYD 41 41
 FT CARBOHYD 96 96
 FT CARBOHYD 107 107
 FT CARBOHYD 114 114
 FT SEQUENCE 155 AA; 17296 MW; F8264E7A32B20D87 CRC64;
 SO
 Query Match Best Local Similarity 90.0%; Score 27; DB 1; Length 155;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YSQCR 5
 DB 115 YSQCR 119
 RESULT 4
 ECP_RAT STANDARD; PRT; 155 AA.
 ID ECP_RAT
 AC P70709;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Eosinophil cationic protein precursor (EC 3.1.27.-) (ECP)
 DE (Ribonuclease 3) (Rnase 3).
 GN RNASE3 OR RNS3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Bone marrow;
 RA MEDLINE=97236431; PubMed=9116043;
 RA Nitton T., Hirakata M., Mue S., Ohuchi K.;
 RT Identification of cDNA encoding rat eosinophil cationic
 RT protein/eosinophil-associated ribonuclease.*;
 RL Blochim. Biophys. Acta 1351:42-46(1997).
 CC -1- FUNCTION: CYTOTOXIN AND HELMINTHOTOXIN WITH RIBONUCLEASE ACTIVITY.
 CC POSSESSES A WIDE VARIETY OF BIOLOGICAL ACTIVITIES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MATRIX OF EOSINOPHIL'S LARGE SPECIFIC
 CC GRANULE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC
CC EMBL; D88586; BAI13648.1; -.
CC HSSP; P00656; 2RNS.
CC InterPro: IPR001427; RNaseA.
CC Pfam; PF00074; RNaseA; 1.
CC ProDom; PD000535; RNaseA; 1.
CC SMART; SM00092; RNase_Pc; 1.
CC PROSITE; PS00127; RNASE_PANCREATIC; 1.
CC Hydrolyase; Nuclease; Endonuclease; Eosinophil; Toxin; Glycoprotein;
CC Signal.
CC FT SIGNAL 1 25 POTENTIAL.
CC FT CHAIN 26 155 EOSINOPHIL CATIONIC PROTEIN.
CC FT ACT_SITE 38 38 BY SIMILARITY.
CC FT ACT_SITE 62 62 BY SIMILARITY.
CC FT ACT_SITE 150 150 BY SIMILARITY.
CC FT DISULFID 47 106 BY SIMILARITY.
CC FT DISULFID 61 118 BY SIMILARITY.
CC FT DISULFID 79 133 BY SIMILARITY.
CC FT DISULFID 86 94 BY SIMILARITY.
CC SQ SEQUENCE 155 AA; 18007 MW; F190864C34EA22C7 CRC64;
Query Match 90.0%; Score 27; DB 1; Length 155;
Best Local Similarity 80.0%; Pred. No. 55;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
Db 115 YTOCR 119

RESULT 5
ECP2_MOUSE STANDARD; PRT; 156 AA.
ID ECP2_MOUSE
AC P97425;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Eosinophil cationic protein 2 precursor (EC 3.1.27.-) (ECP 2) (Ribonuclease 3-2) (RNase 3-2) (Eosinophil secondary granule ribonuclease-2) (EAR-2).
GN EAR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=97057249; PubMed=8901588;
RA Larson K.A., Olson E.V., Madden B.J., Gleich G.J., Lee N.A., Lee J.J.;
RT "Two highly homologous ribonuclease genes expressed in mouse eosinophils identify a larger subgroup of the mammalian ribonuclease superfamily";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12370-12375(1996).
CC -1- FUNCTION: CYTOTOXIN AND HELMINTHOTOXIN WITH RIBONUCLEASE ACTIVITY.
CC POSSESSES A WIDE VARIETY OF BIOLOGICAL ACTIVITIES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MATRIX OF EOSINOPHIL'S LARGE SPECIFIC GRANULE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

HSSP; P00656; 2RNS.
MCD; MGI:108020; Ear2.
InterPro: IPR001427; RNaseA.
Pfam; PF00074; RNaseA; 1.
PRINTS; PR00794; RIBONUCLEASE.
ProDom; PD000535; RNaseA; 1.
SMART; SM00092; RNase_Pc; 1.
PROSITE; PS00127; RNASE_PANCREATIC; 1.
Hydrolyase; Nuclease; Endonuclease; Eosinophil; Toxin; Glycoprotein;
Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 156 EOSINOPHIL CATIONIC PROTEIN 2.
FT ACT_SITE 38 38 BY SIMILARITY.
FT ACT_SITE 62 62 BY SIMILARITY.
FT ACT_SITE 151 151 BY SIMILARITY.
FT DISULFID 47 106 BY SIMILARITY.
FT DISULFID 61 119 BY SIMILARITY.
FT DISULFID 79 134 BY SIMILARITY.
FT CARBOHYD 86 94 BY SIMILARITY.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 156 AA; 17620 MW; A158C1133091053 CRC64;
Query Match 90.0%; Score 27; DB 1; Length 156;
Best Local Similarity 80.0%; Pred. No. 55;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
Db 116 YTOCR 120

RESULT 6
ECP3_MOUSE STANDARD; PRT; 156 AA.
ID ECP3_MOUSE
AC O35290;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eosinophil cationic-type ribonuclease 3 precursor (MR-3).
GN EAR3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=97478550; PubMed=9336452;
RA Batten D., Dyer K.D., Domachowski J.B., Rosenberg H.F.;
RT "Molecular cloning of four novel murine ribonuclease genes: unusual expansion within the ribonuclease A gene family";
RL Nucleic Acids Res. 25:4235-4239(1997).
CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF017258; AAC53489.1; -.
CC HSSP; P00656; 2RNS.
CC MCD; MGI:1858237; Ear3.
CC InterPro: IPR001427; RNaseA.
CC Pfam; PF00074; RNaseA; 1.
CC PRINTS; PR00794; RIBONUCLEASE.
CC ProDom; PD000535; RNaseA; 1.
CC SMART; SM00092; RNase_Pc; 1.
CC PROSITE; PS00127; RNASE_PANCREATIC; 1.

KW Hydrolyase; Nuclease; Endonuclease; Glycoprotein; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 156 ECOSINOPHIL CATIONIC-TYPE RIBONUCLEASE 3.
 FT ACT_SITE 38 38 BY SIMILARITY.
 FT ACT_SITE 62 62 BY SIMILARITY.
 FT ACT_SITE 151 151 BY SIMILARITY.
 FT DISULFID 47 106 BY SIMILARITY.
 FT DISULFID 61 119 BY SIMILARITY.
 FT DISULFID 79 134 BY SIMILARITY.
 FT DISULFID 86 94 BY SIMILARITY.
 FT CARBOHYD 41 41 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 89 89 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 107 107 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SQ SEQUENCE 156 AA: 17744 MW: 17744 C9C2B6CEB5DA06C CRC64;

Query Match 90.0%; Score 27; DB 1; Length 156;
 Best Local Similarity 80.0%; Pred. No. 55;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YSQR 5
 DB 116 YTCR 120

RESULT 7

YB28_MYCTU STANDARD; PRT: 451 AA.
 AC 006580;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein Rv1128c.
 GN Rv1128c OR M1160 OR M12268.17C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37R;
 RX MEDLINE=9825987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekla F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Uitterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE Rv1128c/1148c/1588c/1702c/1945/3466
 CC FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.

CC EMBL; Z95985; CAB09039.1;
 DR EMBL; AE006995; AAK45417.1;
 DR TIGR; MT1160;
 DR Tuberculist; Rv1128c;
 DR InterPro; IPR003870; DUF222.
 DR InterPro; IPR003615; HNHunc.
 DR Pfam; PF02720; DUF222; 1.
 DR SMART; SM00507; HNHc; 1.
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 270 E -> G (IN REF. 2).
 FT CONFLICT 421 MISSING (IN REF. 2).
 SQ SEQUENCE 451 AA: 49274 MW: 469774E5CE148CDD CRC64;

Query Match 90.0%; Score 27; DB 1; Length 451;
 Best Local Similarity 80.0%; Pred. No. 1,3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YSQR 5
 DB 383 YACR 387

RESULT 8

AKA8_HUMAN STANDARD; PRT: 692 AA.
 AC 043823;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE A-kinase anchor protein 8 (A-kinase anchor protein 95 kDa) (AKAP 95).
 GN AKA8 OR AKAP95.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cerebellum, and Testis;
 RX MEDLINE=98142017; PubMed=947338;
 RA Elde T., Coghlan V., Oestavik S., Holtsve C., Solberg R.,
 RA Skeelheg B.S., Lamb N.C., Langeberg L., Fernandez A., Scott J.D.,
 RA Janssen T., Tasken K.;
 RT "Molecular cloning, chromosomal localization, and cell cycle-dependent
 RT subcellular distribution of the A-kinase anchoring protein, AKAP95,"
 RL Exp. Cell Res. 238:305-316(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
 RA Burkhart-Schultz K., Gordon L., Dias J., Ramirez M., Stiliwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
 RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Altix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas R., Quan G., Krommiller B.,
 RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrino A.V.;
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ANCHORING PROTEIN THAT MEDIATES THE SUBCELLULAR
 CC COMPARTMENTATION OF CAMP-DEPENDENT PROTEIN KINASE (PKA TYPE II).
 CC -1- SUBUNIT: BINDS TO DIMERIC RII-ALPHA REGULATORY SUBUNIT OF PKA
 CC DURING MITOSIS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
 CC REDISTRIBUTED AND DETACHED FROM CONDENSED CHROMATIN DURING
 CC MITOSIS.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, LIVER, SKELETAL
 CC MUSCLE, KIDNEY AND PANCREAS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Y11997: CAA2722.1: -
DR EMBL: AC005785: AAC62838.1: -
DR Gene: HGNC:378: AKAP8.
DR MIM: 604692: -
DR InterPro: IPR000822: znf_C2H2.
DR SMART: SM00355: znf_C2H2.1.
KW Nuclear protein; Zinc-finger; DNA-binding.
FT DOMAIN 107 118 POLY-GLY.
FT DOMAIN 368 377 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT ZN_FING 392 414 C2H2-TYPE.
FT ZN_FING 481 504 C2H2-TYPE.
FT DOMAIN 572 589 RII-BINDING (BY SIMILARITY).
SQ SEQUENCE 692 AA: 76108 MW: CACD5F014FDP94B66 CRC64;

Query Match 90.0%; Score 27; DB 1; Length 692;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSQR 5
DB 170 YSECR 174

RESULT 9
POLN_HEYBU STANDARD: PRT: 1691 AA.
ID POLN_HEYBU
AC P29324:
DT 01-DEC-1993 (Rel. 27, Created)
DT 01-DEC-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-structural polypeptide [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Helicase].
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Hung C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Fernandez G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV)."
RL Virology 191:550-558(1992).
RN [2]
RN SEQUENCE OF 965-1691 FROM N.A.
RX MEDLINE=92271462; PubMed=158964;
RA Fry K.E., Tam A.W., Smith M.M., Kim J.P., Luk K.C., Young L.M.,
RA Platak M., Feldman R.A., Yun K.Y., Purdy M.A., McCausland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site."
RL Virus Genes 6:173-185(1992).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- MISCELLANEOUS: HEPATITIS E VIRUS IS THE MAJOR CAUSATIVE AGENT OF
CC ENTERICALLY TRANSMITTED NON-A, NON-B HEPATITIS (ET-NANBH).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M74506: AAA5730.1: -
DR PIR: A44212: A44212.
DR MEROPS: CA1.001: -.

DR InterPro: IPR002589; A1PP.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR002588; V_methyltransf.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
DR Pfam: PF01660; Vmethyltransf; 1.
DR Pfam: PF01661; A1PP; 1.
DR SMART: SM00506; A1PP; 1.
KW Polypeptide; Transferase; RNA-directed RNA polymerase; Helicase;
KW ATP-binding.
FT NP_BIND 973 980 ATP (POTENTIAL).
SQ SEQUENCE 1691 AA: 185224 MW: DB3F0B2C913F871B CRC64;

Query Match 90.0%; Score 27; DB 1; Length 1691;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSQR 5
DB 431 YAQCR 435

RESULT 10
POLN_HEYBU STANDARD: PRT: 1693 AA.
ID POLN_HEYBU
AC P29324:
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-structural polypeptide [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Helicase].
OS Hepatitis E virus (strain Burma) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome."
RL Virology 185:120-131(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- MISCELLANEOUS: HEPATITIS E VIRUS IS THE MAJOR CAUSATIVE AGENT OF
CC ENTERICALLY TRANSMITTED NON-A, NON-B HEPATITIS (ET-NANBH).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M73218: AAA5734.1: -
DR PIR: A40778; MNMWE.
DR MEROPS: CA1.001: -
DR InterPro: IPR002589; A1PP.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR002588; V_methyltransf.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
DR Pfam: PF01660; Vmethyltransf; 1.
DR Pfam: PF01661; A1PP; 1.
DR SMART: SM00506; A1PP; 1.
KW Polypeptide; Transferase; RNA-directed RNA polymerase; Helicase;
KW ATP-binding.
FT NP_BIND 975 982 ATP (POTENTIAL).
FT NP_BIND 975 982 ATP (POTENTIAL).

SQ SEQUENCE 1693 AA; 185191 MW; 2F355E46E9ED219B CRC64;
 Query Match 90.0%; Score 27; DB 1; Length 1693;
 Best Local Similarity 80.0%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 YSOCR 5
 DB 431 YAOGR 435
 RESULT 11
 ID POLN_HEVPA STANDARD; PRT: 1693 AA.
 AC 004610;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Non-structural polypeptide [Contains: RNA-directed RNA polymerase
 (EC 2.7.7.48); Helicase].
 OS Hepatitis E virus (strain Myanmar) (HEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=31769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93227573; PubMed=8470371;
 RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
 RA Rikhsisa T., Wain K.;
 RT "Sequence and gene structure of the hepatitis E virus isolated from
 Myanmar.";
 RL Virus News 7:95-109(1993).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
 (RNA)(N).
 CC -1- MISCELLANEOUS: HEPATITIS E VIRUS IS THE MAJOR CAUSATIVE AGENT OF
 ENTERICALLY TRANSMITTED NON-A, NON-B HEPATITIS (ET-NANBH).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10330; BAA01172.1; -
 DR MEROPS: C41.001; -
 DR InterPro: IPR002589; Alpp.
 DR InterPro: IPR001230; Prenyl_site.
 DR InterPro: IPR001788; RNA_dep_RNApol2.
 DR InterPro: IPR002588; V_methyltransf.
 DR InterPro: IPR000606; Viral_helicase1.
 DR Pfam: PF00978; RNA_dep_RNApol2; 1.
 DR Pfam: PF01443; Viral_helicase1; 1.
 DR Pfam: PF01660; Vmethyltransf; 1.
 DR SMART: SM00506; Alpp; 1.
 DR SMART: SM00506; Alpp; 1.
 DR Polyprotein; Transferase; RNA-directed RNA polymerase; Helicase;
 KW ATP-binding.
 FT NP-BIND 975
 FT SEQUENCE 1693 AA; 185215 MW; AAB4C9140A7E21EA CRC64;
 SQ
 Query Match 90.0%; Score 27; DB 1; Length 1693;
 Best Local Similarity 80.0%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 YSOCR 5
 DB 431 YAOGR 435
 RESULT 12
 POLN_HEVPA

ID POLN_HEVPA STANDARD; PRT: 1693 AA.
 AC P33424;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Non-structural polypeptide [Contains: RNA-directed RNA polymerase
 (EC 2.7.7.48); Helicase].
 OS Hepatitis E virus (strain Pakistan) (HEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=33774;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92115700; PubMed=1731327;
 RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
 RA Malik I.A., Iqbal M., Purcell R.H.;
 RT "Characterization of a prototype strain of hepatitis E virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
 (RNA)(N).
 CC -1- MISCELLANEOUS: HEPATITIS E VIRUS IS THE MAJOR CAUSATIVE AGENT OF
 ENTERICALLY TRANSMITTED NON-A, NON-B HEPATITIS (ET-NANBH).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M80581; AAA45725.1; -
 DR MEROPS: C41.001; -
 DR InterPro: IPR002589; Alpp.
 DR InterPro: IPR001230; Prenyl_site.
 DR InterPro: IPR001788; RNA_dep_RNApol2.
 DR InterPro: IPR002588; V_methyltransf.
 DR InterPro: IPR000606; Viral_helicase1.
 DR Pfam: PF00978; RNA_dep_RNApol2; 1.
 DR Pfam: PF01443; Viral_helicase1; 1.
 DR Pfam: PF01660; Vmethyltransf; 1.
 DR SMART: SM00506; Alpp; 1.
 DR SMART: SM00506; Alpp; 1.
 DR Polyprotein; Transferase; RNA-directed RNA polymerase; Helicase;
 KW ATP-binding.
 FT NP-BIND 975
 FT SEQUENCE 1693 AA; 185149 MW; 5A0F03FBF1B9958 CRC64;
 SQ
 Query Match 90.0%; Score 27; DB 1; Length 1693;
 Best Local Similarity 80.0%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 YSOCR 5
 DB 431 YAOGR 435
 RESULT 13
 ID RM16_MAIZE STANDARD; PRT: 185 AA.
 AC P27927;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Mitochondrial 60S ribosomal protein L16.
 GN RPL16.
 OS Zea mays (Maize).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoidae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-91216097; PubMed-1708720;
 RA Hunt M.D., Newton K.J.;
 RT "The NCS1 mutation: genetic evidence for the expression of ribosomal
 RT protein genes in 26a ways mitochondria.";
 RL EMBL J. 10:1045-1052(1991).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X57445; CAA04691.1; -
 DR EMBL; X57445; CAA04691.1; -
 DR PIR; S15026; R52M6M.
 DR MaltzDB; 69578; -
 DR InterPro: IPR000114; Ribosomal_L16.
 DR Pfam; PF00252; Ribosomal_L16; 1.
 DR PRINTS; PR00060; RIBOSOMAL_L16.
 DR TIGRFAMS; TIGR01164; rplp_bact; 1.
 DR PROSITE; PS00586; RIBOSOMAL_L16_1; 1.
 DR PROSITE; PS00701; RIBOSOMAL_L16_2; 1.
 KW Ribosomal protein; Mitochondrion.
 SQ SEQUENCE 185 AA; 21033 MW; F9CD590F773506A5 CRC64;

Query Match 86.7%; Score 26; DB 1; Length 185;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSCOR 5
 Db 53 YSKCR 57

RESULT 14
 RML6_ORYSA STANDARD; PRT; 185 AA.
 ID RML6_ORYSA
 AC P46801;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Mitochondrial 60S ribosomal protein L16.
 GN RPL16.
 OS Oryza sativa (Rice).
 OG Mitochondrion.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Eriatloideae; Oryzaceae; Oryza.
 CC NCBI_TaxID=4530;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Nipponbare / Japonica; TISSUE=Leaf;
 RX MEDLINE-95308541; PubMed-7788722;
 RA Nakazono M., Iitadani H., Wakasugi T., Tsutsumi N., Sugitara M.,
 RA Hirai A.;
 RT "The rps3-rpl16-nad3-rps12 gene cluster in rice mitochondrial DNA is
 RT transcribed from alternative promoters.";
 RL Curr. Genet. 27:184-189(1995).
 CC [2]
 CC SEQUENCE OF 42-185 FROM N.A.
 CC STRAIN=cv. Nipponbare / Japonica;
 CC MEDLINE-95211382; PubMed-7545979;
 RA Iitadani H., Wakasugi T., Sugita M., Sugitara M., Nakazono M.,
 RA Hirai A.;
 RT "Nucleotide sequence of a 28-kbp portion of rice mitochondrial DNA:
 RT the existence of many sequences that correspond to parts of
 RT mitochondrial genes in intergenic regions.";
 RL Plant Cell Physiol. 35:1239-1244(1994).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.

CC -1- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D21251; BAA04793.1; -
 DR EMBL; D21251; BAA04793.1; ALT-TERM.
 DR InterPro: IPR000114; Ribosomal_L16.
 DR Pfam; PF00252; Ribosomal_L16; 1.
 DR PRINTS; PR00060; RIBOSOMAL_L16.
 DR TIGRFAMS; TIGR01164; rplp_bact; 1.
 DR PROSITE; PS00586; RIBOSOMAL_L16_1; 1.
 DR PROSITE; PS00701; RIBOSOMAL_L16_2; 1.
 KW Ribosomal protein; Mitochondrion.
 SQ SEQUENCE 185 AA; 20961 MW; 5FA75CDEFC4C8539 CRC64;

Query Match 86.7%; Score 26; DB 1; Length 185;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSCOR 5
 Db 53 YSKCR 57

RESULT 15
 Y460_PYRHO STANDARD; PRT; 188 AA.
 ID Y460_PYRHO
 AC 058215;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein PH0460.
 GN PH0460.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Pyrococcus.
 CC NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE-98344137; PubMed-9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 CC -1- SIMILARITY: BELONGS TO THE CDP-ALCOHOL PHOSPHATIDYLTRANSFERASE
 CC CLASS-1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP000002; BAA29546.1; -
 DR InterPro: IPR000462; CDP-OH_P_transf.
 DR Pfam; PF01066; CDP-OH_P_transf; 1.
 DR PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; 1.
 KW Hypothetical protein; Transferase; Complete proteome.
 SQ SEQUENCE 188 AA; 20418 MW; D7FE6234216D3214 CRC64;

Query Match 86.7%; Score 26; DB 1; Length 188;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOGR 5
 11:11
 Db 123 YSRGR 127

Search completed: July 9, 2003, 12:00:44
 Job time : 8.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 9, 2003, 11:57:21 ; Search time 17.75 Seconds

(without alignments)
58.041 Million cell updates/sec

Title: US-09-851-058-2

Perfect score: 30

Sequence: 1 YSOCR 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_podent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	30	100.0	87 3 Q8TGU2	Q8TGU2 saccharomyc
2	30	100.0	134 11 Q9JKI1	Q9JKI1 mus pahari
3	30	100.0	155 11 Q9JRH6	Q9JRH6 mus caroli
4	30	100.0	155 11 Q9JRG8	Q9JRG8 mus caroli
5	30	100.0	155 11 Q9JRG2	Q9JRG2 mus caroli
6	30	100.0	221 5 Q62157	Q62157 ctenorhabdi
7	30	100.0	295 10 Q9FTU3	Q9FTU3 oryza sativ
8	30	100.0	341 16 Q9PFI1	Q9PFI1 xylella fas
9	30	100.0	367 2 Q9AGI1	Q9AGI1 pseudomonas
10	30	100.0	555 10 Q94E22	Q94E22 arabidopsis
11	30	100.0	555 10 Q80612	Q80612 arabidopsis
12	27	90.0	78 12 Q98440	Q98440 paramedum
13	27	90.0	153 11 Q9R130	Q9R130 rattus norv
14	27	90.0	153 11 Q9JKI7	Q9JKI7 mus saxicol
15	27	90.0	155 11 Q9R134	Q9R134 rattus norv
16	27	90.0	155 11 Q9R132	Q9R132 rattus norv

17	27	90.0	155 11 Q9R131	Q9R131 rattus norv
18	27	90.0	155 11 Q9R127	Q9R127 rattus norv
19	27	90.0	155 11 Q9R125	Q9R125 mus musculu
20	27	90.0	155 11 Q9JKI6	Q9JKI6 mus saxicol
21	27	90.0	155 11 Q9JKI4	Q9JKI4 mus saxicol
22	27	90.0	155 11 Q9JKI2	Q9JKI2 mus saxicol
23	27	90.0	155 11 Q9JKI0	Q9JKI0 mus pahari
24	27	90.0	155 11 Q9JKH9	Q9JKH9 mus pahari
25	27	90.0	155 11 Q9JKH8	Q9JKH8 mus pahari
26	27	90.0	155 11 Q9JKH3	Q9JKH3 mus caroli
27	27	90.0	155 11 Q9JKG3	Q9JKG3 mus caroli
28	27	90.0	155 11 Q9R126	Q9R126 mus musculu
29	27	90.0	155 11 Q923L7	Q923L7 mus musculu
30	27	90.0	156 11 Q9R124	Q9R124 mus musculu
31	27	90.0	156 11 Q9JKH7	Q9JKH7 mus caroli
32	27	90.0	156 11 Q9JKH5	Q9JKH5 mus caroli
33	27	90.0	156 11 Q9JKH4	Q9JKH4 mus caroli
34	27	90.0	156 11 Q9JKH2	Q9JKH2 mus caroli
35	27	90.0	156 11 Q9JKH1	Q9JKH1 mus caroli
36	27	90.0	156 11 Q9JKH0	Q9JKH0 mus caroli
37	27	90.0	156 11 Q9JKG9	Q9JKG9 mus caroli
38	27	90.0	156 11 Q9JKG7	Q9JKG7 mus caroli
39	27	90.0	156 11 Q9JKG6	Q9JKG6 mus caroli
40	27	90.0	156 11 Q923L6	Q923L6 mus musculu
41	27	90.0	156 11 Q91207	Q91207 mus musculu
42	27	90.0	156 11 Q8VHS0	Q8VHS0 mus musculu
43	27	90.0	157 11 Q9JKJ8	Q9JKJ8 citellus
44	27	90.0	275 12 Q9PXB1	Q9PXB1 periplaneta
45	27	90.0	310 8 Q63637	Q63637 balclutha 1

ALIGNMENTS

RESULT 1
Q8TGU2 ID Q8TGU2 PRELIMINARY; PRT; 87 AA.
AC Q8TGU2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 10.0 kDa protein.
GN YFRO32C-B.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Saccharomycetes.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid:4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21624570; Pubmed-11753363;
RA Kumar A., Harrison P.M., Cheung K.H., Lan N., Echols N., Bertone P., Miller P., Gerstein M.B., Snyder M.;
RT "An integrated approach for finding overlooked genes in yeast.";
RL Nat. Biotechnol. 20:58-63(2002).
DR EMBL; AF479893; AAL79206.1; -;
KW Hypothetical protein.
SO SEQUENCE 87 AA; 9953 MW; 799BF7BD1A45D427 CRC64;

Query Match 100.0%; Score 30; DB 3; Length 87;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSOCR 5
DB 12 YSOCR 16

RESULT 2
Q9JKI1 ID Q9JKI1 PRELIMINARY; PRT; 154 AA.
AC Q9JKI1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Eosinophil-associated ribonuclease 2.
GN EAR2.
OS Mus pahari (Shrew mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10093;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20243759; PubMed=10758160;
RA Zhang J., Dyer K.D., Rosenberg H.F.;
RT "Evolution of the rodent eosinophil-associated ribonuclease gene
family by rapid gene sorting and positive selection."
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
DR EMBL: AF238402; AAF67702.1; -
DR HSSP: P00656; ILSO.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA.1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR ProDom: PD000535; RNaseA.1.
DR SMART: SM00092; RNase_Pc.1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN.1.
SQ SEQUENCE 154 AA; 17927 MW; 08F002D25B461DDE CRC64;

Query Match 100.0%; Score 30; DB 11; Length 154;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSQR 5
DB 114 YSQR 118

RESULT 3
Q9JKH6
ID Q9JKH6 PRELIMINARY; PRT; 155 AA.
AC Q9JKH6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Eosinophil-associated ribonuclease 2.
GN EAR2.
OS Mus caroli (Wild mouse) (Ricefield mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10089;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20243759; PubMed=10758160;
RA Zhang J., Dyer K.D., Rosenberg H.F.;
RT "Evolution of the rodent eosinophil-associated ribonuclease gene
family by rapid gene sorting and positive selection."
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
DR EMBL: AF238407; AAF67707.1; -
DR HSSP: P00656; ILSO.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA.1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR ProDom: PD000535; RNaseA.1.
DR SMART: SM00092; RNase_Pc.1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN.1.
SQ SEQUENCE 155 AA; 17257 MW; 689F290CF973328C CRC64;

Query Match 100.0%; Score 30; DB 11; Length 155;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSQR 5
DB 115 YSQR 119

RESULT 4

Q9JKG8
ID Q9JKG8 PRELIMINARY; PRT; 155 AA.
AC Q9JKG8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Eosinophil-associated ribonuclease 21.
GN EAR21.
OS Mus caroli (Wild mouse) (Ricefield mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10089;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20243759; PubMed=10758160;
RA Zhang J., Dyer K.D., Rosenberg H.F.;
RT "Evolution of the rodent eosinophil-associated ribonuclease gene
family by rapid gene sorting and positive selection."
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
DR EMBL: AF238416; AAF67715.1; -
DR HSSP: P00656; ILSO.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA.1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR ProDom: PD000535; RNaseA.1.
DR SMART: SM00092; RNase_Pc.1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN.1.
SQ SEQUENCE 155 AA; 17420 MW; 56B351AA2B784551 CRC64;

Query Match 100.0%; Score 30; DB 11; Length 155;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSQR 5
DB 115 YSQR 119

RESULT 5
Q9JKG2
ID Q9JKG2 PRELIMINARY; PRT; 155 AA.
AC Q9JKG2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Eosinophil-associated ribonuclease 33.
GN EAR33.
OS Mus caroli (Wild mouse) (Ricefield mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10089;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20243759; PubMed=10758160;
RA Zhang J., Dyer K.D., Rosenberg H.F.;
RT "Evolution of the rodent eosinophil-associated ribonuclease gene
family by rapid gene sorting and positive selection."
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
DR EMBL: AF238422; AAF67721.1; -
DR HSSP: P00656; ILSO.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA.1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR ProDom: PD000535; RNaseA.1.
DR SMART: SM00092; RNase_Pc.1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN.1.
SQ SEQUENCE 155 AA; 17357 MW; C5251FAB85197A42 CRC64;

Query Match 100.0%; Score 30; DB 11; Length 155;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YSQR 5

Db 11111
115 YSOCR 119

RESULT 6

062157 ID 062157 PRELIMINARY; PRT; 221 AA.
AC 062157;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE F1D11.8 protein.
GN F1D11.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Mortimore B.J., Basham V.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; Pubmed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z81500; CAB04099.1; -;
SQ SEQUENCE 221 AA; 25546 MW; 221F6EDF6844912F CRC64;

Query Match 100.0%; Score 30; DB 5; Length 221;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
11111
Db 107 YSOCR 111

RESULT 7

09FTU3 ID 09FTU3 PRELIMINARY; PRT; 295 AA.
AC 09FTU3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Putative branched-chain amino acid aminotransferase protein.
GN OSJNB0086P08.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatloidea; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:OSJNB0086P08.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002855; BAB17203.1; -;
DR InterPro: IPR001544; AminoTtran-4.
DR Pfam; PF01063; aminoTtran_4; 1.
DR ProDom; PD001961; AminoTtran_4; 1.
KW AminoTransferase; Transferase.
SQ SEQUENCE 295 AA; 31705 MW; 4C26272EE9410D9 CRC64;

Query Match 100.0%; Score 30; DB 10; Length 295;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5

Db 11111
120 YSOCR 124

RESULT 8

09PFV1 ID 09PFV1 PRELIMINARY; PRT; 341 AA.
AC 09PFV1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein Xf0556.
GN Xf0556.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; Pubmed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
Coutinho L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
Fachinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
Ho P.L., Honelisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
Krieger J.E., Kuramae E.E., Laigret F., Lambis M.R., Leite L.C.C.,
Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
Machado M.A., Madeira A.M.B.N., Martins E.M.F., Marino C.L.,
Marques M.V., Martins E.A.L., Martins E.M.F., Matsushima A.T.,
Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
Peleto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
Zago M.A., Zatz M., Zeldanis J., Zetubai J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE003903; AAF83366.1; -;
DR InterPro: IPR004881; DUF258.
DR Pfam; PF03193; DUF258; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 341 AA; 37432 MW; 32AC71ECC7D0DF3A CRC64;

Query Match 100.0%; Score 30; DB 16; Length 341;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOCR 5
11111
Db 34 YSOCR 38

RESULT 9

09AG11 ID 09AG11 PRELIMINARY; PRT; 367 AA.
AC 09AG11;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Hypothetical 40.8 kDa protein.
OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21374446; PubMed=11481492;
 RT "A genomic island in *Pseudomonas aeruginosa* carries the determinants
 of flagellin glycosylation."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9342-9347(2001).
 DR EMBL: AF332547; AAK15337.1;
 DR InterPro: IPR000051; SAM_bind.
 KW Hypothetical protein.
 SO SEQUENCE 367 AA; 40811 MW; D83BE4716FCB2A7B CRC64;

Query Match 100.0%; Score 30; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSQCR 5
 Db 111 YSQCR 115

RESULT 10
 ID 094E22 PRELIMINARY; PRT; 555 AA.
 AC 094E22;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative nucleoside triphosphatase.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
 RA Palm C.D., Bower L., Jones T., Banh J., Carinci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kaniya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinzaki K.,
 RA Becker J., Theologis A., Davis R.W.,
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A8387012; AAK62457.1;
 DR InterPro: IPR000407; GDA1-CD39; NTase.
 DR Pfam: PF01150; GDA1-CD39; 1.
 SO SEQUENCE 555 AA; 61318 MW; 0D1A9FDD4C436962 CRC64;

Query Match 100.0%; Score 30; DB 10; Length 555;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSQCR 5
 Db 349 YSQCR 353

RESULT 11
 ID 080612 PRELIMINARY; PRT; 555 AA.
 AC 080612;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative nucleoside triphosphatase.
 GN AT3G02970.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC004138; AAC32915.2;
 SO SEQUENCE 555 AA; 61287 MW; 10DBE8662132927 CRC64;

Query Match 100.0%; Score 30; DB 10; Length 555;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSQCR 5
 Db 349 YSQCR 353

RESULT 12
 ID 098440 PRELIMINARY; PRT; 78 AA.
 AC 098440;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE A388R protein.
 GN A388R.
 OS Paramoecium bursaria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 OX NCBI_TaxID=10506;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96400190; PubMed=8806566;
 RA Kutish G.F., Li Y., Lu Z., Furuta M., Rock D.L., Van Etten J.L.;
 RT Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map
 RT positions 182 to 258."
 RL Virology 223:303-317(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20013326; PubMed=10544099;
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
 RA Liase A.D., Nickerson K.W., Van Etten J.L.;
 RT "Chlorella virus PBCV-1 encodes a functional homospemidine
 RT synthase."
 RL Virology 263:254-262(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20478054; PubMed=11021991;
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
 RT PBCV-1."
 RL Virology 276:27-36(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.

RA Van Etten J.L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Gurnon J.R., Graves M.V., Van Etten J.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U42580; AAC96756.1; -
SQ SEQUENCE 78 AA; 9346 MW; 234659C39B183E7F CRC64;

Query Match 90.0%; Score 27; DB 12; Length 78;
Best Local Similarity 80.0%; Pred. No. 40;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
Db 27 YAOGR 31

RESULT 13
OGR130 PRELIMINARY; PRT; 153 AA.
ID OGR130;
AC OGR130;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Ribonuclease 7 precursor (Fragment).
GN R7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20063779; PubMed-10594173;
RA Singhania N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RA Domachowski J.B., Rosenberg H.F.;
RT "Rapid evolution of the ribonuclease A superfamily: adaptive expansion
of independent gene clusters in rats and mice.";
RL J. Mol. Evol. 49:721-728(1999).
DR EMBL: AF171645; AAD51665.1; -
DR HSSP: P00656; 2RNS.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA.
DR PRINTS: PR00794; RIBONUCLEASE.
DR PRODOM: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase_Pc; 1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 >153 RIBONUCLEASE 7.
FT NON_TER 153 153
SQ SEQUENCE 153 AA; 17679 MW; B5878251BA734D6B CRC64;

Query Match 90.0%; Score 27; DB 11; Length 153;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
Db 113 YTOGR 117

RESULT 14
O9JK17 PRELIMINARY; PRT; 153 AA.

AC O9JK17;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Eosinophil-associated ribonuclease 4.
GN EAR4.
OS Mus saxicola (Spiny mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20243759; PubMed-10758160;
RA Zhang J., Dyer K.D., Rosenberg H.F.;
RT "Evolution of the rodent eosinophil-associated ribonuclease gene
family by rapid gene sorting and positive selection.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
DR EMBL: AF238396; AAF67696.1; -
DR HSSP: P00656; 2RNS.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA.
DR PRINTS: PR00794; RIBONUCLEASE.
DR PRODOM: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase_Pc; 1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
SQ SEQUENCE 153 AA; 17042 MW; CFF9982F0A6FAF296 CRC64;

Query Match 90.0%; Score 27; DB 11; Length 153;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YSOGR 5
Db 115 YTOGR 119

RESULT 15
OGR134 PRELIMINARY; PRT; 155 AA.
ID OGR134;
AC OGR134;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Ribonuclease 1 precursor (Fragment).
GN R1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20063779; PubMed-10594173;
RA Singhania N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RA Domachowski J.B., Rosenberg H.F.;
RT "Rapid evolution of the ribonuclease A superfamily: adaptive expansion
of independent gene clusters in rats and mice.";
RL J. Mol. Evol. 49:721-728(1999).
DR EMBL: AF171641; AAD51661.1; -
DR HSSP: P00656; 2RNS.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA.
DR PRINTS: PR00794; RIBONUCLEASE.
DR PRODOM: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase_Pc; 1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 >155 RIBONUCLEASE 1.
FT NON_TER 155 155
SQ SEQUENCE 155 AA; 18119 MW; 16BAA6503260CEB CRC64;

Query Match 90.0%; Score 27; DB 11; Length 155;
Best Local Similarity 80.0%; Pred. No. 75;

Matches	4;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	YSOCR	5						
		1:111							
Db	115	YTOCR	119						

Search completed: July 9, 2003, 12:02:03
Job time: 19.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:55:35 ; Search time 22 Seconds
(Without alignments)
30.284 Million cell updates/sec

Title: US-09-851-058-3
Perfect score: 31
Sequence: 1 YTCOR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /Geneseq_101002:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	100.0	39	21	AAAG12153	Arabidopsis thaliana
2	31	100.0	134	19	AAAG5178	Rat eosinophil cat
3	31	100.0	155	19	AAAG5177	Rat eosinophil cat
4	31	100.0	155	20	AAAG94272	Rat-derived eosino
5	31	100.0	155	20	AAAG94273	Rat eosinophils-de
6	28	90.3	44	23	ABPA2832	Human ovarian anti
7	28	90.3	105	22	AAO10286	Human polypeptide
8	28	90.3	322	22	AAU44616	Protonibacterium
9	28	90.3	435	21	AAAG24041	Arabidopsis thaliana
10	28	90.3	440	21	AAAG24040	Arabidopsis thaliana

11	28	90.3	441	22	ABB61563	Drosophila melanog
12	28	90.3	476	22	AAU29351	Novel mar regulate
13	28	90.3	493	22	ABG25044	Novel human diagno
14	28	90.3	652	22	ABB67974	Drosophila melanog
15	28	90.3	856	20	AAAG89761	C glutamicum prote
16	27	87.1	35	20	AAI26585	Peptide sequence o
17	27	87.1	69	20	AAI18129	Clone 3 of A. thal
18	27	87.1	87	22	AAO08211	Human polypeptide
19	27	87.1	89	21	AAI18577	Arabidopsis thaliana
20	27	87.1	159	22	ABB52750	Escherichia coli p
21	27	87.1	243	21	AAI28425	Rice branched chl
22	27	87.1	277	22	AAI75049	Human colon cancer
23	27	87.1	307	22	AAI16269	Human novel secret
24	27	87.1	337	21	AAI57130	Human prostate can
25	27	87.1	357	22	ABB64777	Drosophila melanog
26	27	87.1	442	21	AAI53178	Maca mulatta thia
27	27	87.1	451	18	AAI18211	Human integrin-11n
28	27	87.1	452	22	AAI84202	Amino acid sequenc
29	27	87.1	452	23	AAU79414	Human integrin-11n
30	27	87.1	452	23	AAU79415	Human integrin-11n
31	27	87.1	452	23	AAU79421	Human integrin-11n
32	27	87.1	499	22	ABG21472	Novel human diagno
33	27	87.1	512	22	ABG04878	Novel human diagno
34	27	87.1	512	22	ABG05959	Novel human diagno
35	27	87.1	692	22	AAAG3713	Human polypeptide,
36	27	87.1	729	22	AAI75590	Human colon cancer
37	27	87.1	1122	20	AAI26579	Murine telomerase
38	27	87.1	1122	23	ABB06711	Mouse telomerase p
39	27	87.1	2174	22	AAAG38652	Human polypeptide
40	27	87.1	2204	22	AAI21474	Novel human diagno
41	27	87.1	2204	22	ABE11768	Human TRAP240 homo
42	27	87.1	2204	22	AAI40438	Human polypeptide
43	26	83.9	42	21	AAI55841	Arabidopsis thaliana
44	26	83.9	48	19	AAI21324	Human semaphorin I
45	26	83.9	50	22	AAU41708	Protonibacterium

ALIGNMENTS

RESULT 1	AAAG12153	standard; Protein; 39 AA.
XX	AAAG12153	
AC	AAAG12153	
XX	AAAG12153	
DT	17-OCT-2000	(first entry)
XX	AAAG12153	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 11153.	
XX	AAAG12153	
KW	Protein identification; signal transduction pathway; metabolic pathway; hydrolisation assay; genetic mapping; gene expression control; promoter; termination sequence.	
KW	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
XX	06-SEP-2000.	
PD	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-0301439.	
XX	25-FEB-1999; 99US-0121825.	
PR	05-MAR-1999; 99US-0123180.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0133048.
PR 04-MAY-1999; 99US-01332407.
PR 05-MAY-1999; 99US-0133484.
PR 06-MAY-1999; 99US-0133485.
PR 06-MAY-1999; 99US-0133486.
PR 07-MAY-1999; 99US-0133487.
PR 11-MAY-1999; 99US-01332863.
PR 14-MAY-1999; 99US-0133256.
PR 14-MAY-1999; 99US-0133218.
PR 14-MAY-1999; 99US-0133219.
PR 14-MAY-1999; 99US-0133221.
PR 14-MAY-1999; 99US-0133370.
PR 18-MAY-1999; 99US-0133376.
PR 19-MAY-1999; 99US-01334941.
PR 20-MAY-1999; 99US-01335124.
PR 21-MAY-1999; 99US-0133533.
PR 24-MAY-1999; 99US-0133629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158359.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

100.0%; Score 31; DB 21; Length 39;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
11111
DB 14 YTOCR 18

RESULT 2

AAW59178 ID AAW59178 standard; Protein: 134 AA.

AC AAW59178;

DT 13-AUG-1998 (first entry)

DE Rat eosinophil cationic protein #2.

KW Eosinophil cationic protein; ECP; cytotoxic; recombinant protein;
treatment; disease; rat.

OS Rattus sp.

PN JP1011777-A.

PD 12-MAY-1998.

PF 25-OCT-1996; 96JP-0284311.

PR 25-OCT-1996; 96JP-0284311.

PA (NIHA) JAPAN ENERGY CORP.

DR WPI: 1998-325891/29.

PT New DNA coding rat-derived eosinophil cationic protein - used, e.g.
to produce recombinant protein for treatment of diseases

PS Claim 1; Page 12; 14pp; Japanese.

CC This sequence is a novel rat-derived eosinophil cationic protein (ECP)
which exhibits cytotoxic activity. The DNA encoding the ECP protein
can be used for the production of recombinant protein which is useful
for the treatment of diseases.

SO Sequence 134 AA;

Query Match 100.0%; Score 31; DB 19; Length 134;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
11111
DB 94 YTOCR 98

RESULT 3

AAW59177 ID AAW59177 standard; Protein: 155 AA.

AC AAW59177;

DT 13-AUG-1998 (first entry)

DE Rat eosinophil cationic protein #1.

KW Eosinophil cationic protein; ECP; cytotoxic; recombinant protein;
treatment; disease; rat.

OS Rattus sp.

PN JP1011777-A.

PD 12-MAY-1998.

PF 25-OCT-1996; 96JP-0284311.

PR 25-OCT-1996; 96JP-0284311.

PA (NIHA) JAPAN ENERGY CORP.

DR WPI: 1998-325891/29.

DR N-PSDB; AAV11934, AAV11935.

PT New DNA coding rat-derived eosinophil cationic protein - used, e.g.
to produce recombinant protein for treatment of diseases

PS Claim 3; Page 9-10; 14pp; Japanese.

CC This sequence is a novel rat-derived eosinophil cationic protein (ECP)
which exhibits cytotoxic activity. The DNA encoding the ECP protein
can be used for the production of recombinant protein which is useful
for the treatment of diseases.

SO Sequence 155 AA;

Query Match 100.0%; Score 31; DB 19; Length 155;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
11111
DB 115 YTOCR 119

RESULT 4

AAW94272 ID AAW94272 standard; Protein: 155 AA.

AC AAW94272;

DT 22-APR-1999 (first entry)

DE Rat-derived eosinophil-derived neurotoxin (EDN).

KW Eosinophil-derived neurotoxin; EDN; eosinophil cationic protein; ECP;
recombinant; rat; precursor.

OS Rattus sp.

PN JP11009280-A.

XX 19-JAN-1999.
 PD 25-JUN-1997; 97JP-0169271.
 PF 25-JUN-1997; 97JP-0169271.
 PR 25-JUN-1997; 97JP-0169271.
 PA (NIHA) JAPAN ENERGY CORP.
 XX WPI; 1999-145897/13.
 DR N-PSDB; AAX05636, AAX05637.
 XX
 PT New rat-derived eosinophil-derived neurotoxin gene - useful for
 preparing recombinant protein
 PS Claim 1; Page 10; 18pp; Japanese.
 CC This represents a rat-derived eosinophil-derived neurotoxin (EDN). The
 CC DNA coding rat-derived EDN or its precursor peptide is useful for the
 CC preparation of a recombinant protein. The base sequence of a precursor
 CC peptide of rat eosinophil-derived eosinophil cationic protein (ECP),
 CC and the base sequence of the precursor peptide of EDN can be determined
 CC by collecting the ECP cDNA and EDN cDNA from eosinophils respectively.
 CC
 SQ Sequence 155 AA;
 Query Match 100.0%; Score 31; DB 20; Length 155;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 YTCR 5
 Db 115 YTCR 119
 RESULT 5
 ID AAW94273 standard; Protein: 155 AA.
 AC AAW94273;
 XX
 DT 22-APR-1999 (first entry)
 DE Rat eosinophils-derived eosinophil cationic protein (ECP).
 DE Eosinophil-derived neurotoxin; EDN; eosinophil cationic protein; ECP;
 KW recombinant; rat; precursor.
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Peptide 59 /note- "indicated as Cys in the sequence listing"
 FT Peptide 99 /note- "indicated as Val in the sequence listing"
 FT Peptide /note- "indicated as Val in the sequence listing"
 PN -JP11009280-A.
 PD 19-JAN-1999.
 XX
 PF 25-JUN-1997; 97JP-0169271.
 PR 25-JUN-1997; 97JP-0169271.
 XX
 PA (NIHA) JAPAN ENERGY CORP.
 XX WPI; 1999-145897/13.
 DR N-PSDB; AAX05638.
 XX
 PT New rat-derived eosinophil-derived neurotoxin gene - useful for
 preparing recombinant protein
 PS Disclosure; Fig 4; 18pp; Japanese.

XX This represents a rat eosinophils-derived eosinophil cationic protein
 CC (ECP). The DNA encoding a rat-derived eosinophil-derived neurotoxin (EDN)
 CC or its precursor peptide is useful for the preparation of a recombinant
 CC protein. The base sequence of a precursor peptide of rat eosinophils-
 CC derived ECP, and the base sequence of the precursor peptide of EDN can
 CC be determined by collecting the ECP cDNA and EDN cDNA from eosinophils
 CC respectively.
 CC
 SQ Sequence 155 AA;
 Query Match 100.0%; Score 31; DB 20; Length 155;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 YTCR 5
 Db 115 YTCR 119
 RESULT 6
 ID ABP42832 standard; Protein: 44 AA.
 AC ABP42832;
 XX
 DT 22-AUG-2002 (first entry)
 DE Human ovarian antigen HPCP68, SEQ ID NO:3964.
 XX
 KW Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX
 OS Homo sapiens.
 XX
 PN WO200200677-A1;
 PD 03-JAN-2002.
 PF 07-JUN-2001; 2001WO-US18569.
 PR 07-JUN-2000; 2000US-209467P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA Blise CE, Rosen CA;
 PI WPI; 2002-147878/19.
 DR N-PSDB; ABQ55909.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11; SEQ ID NO 3964; 2922pp; English.
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related

disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 44 AA:

Query Match 90.3%; Score 28; DB 23; Length 44;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTCR 5
11:11
Db 25 YTECR 29

RESULT 7

AA010288
ID AA010288 standard; Protein: 105 AA.

AC AA010288;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 24180.

Human cytokine; cell proliferation; cell differentiation; gene therapy;

Human vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

tissue growth factor; immunomodulatory; cancer; leukaemia;

nevous system disorders; arthritis; Inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

XX N-PSDB: AA190219.

XX Isolated nucleic acids and polypeptides, useful for preventing

XX diagnosing and treating e.g. leukaemia, inflammation and immune

XX disorders -

XX Claim 20; SEQ ID NO 24180; 1399pp + Sequence Listing; English.

the encoded proteins (AA000010-AA013910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 105 AA:

Query Match 90.3%; Score 28; DB 22; Length 105;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTCR 5
11:11
Db 17 YTECR 21

RESULT 8

AAU44616
ID AAU44616 standard; Protein: 322 AA.

AC AAU44616;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #5512.

SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;

uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71.

XX N-PSDB: AAS59523.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX vaccinating against and diagnosing infections, especially useful for

XX treating acne vulgaris -

XX Example 1; SEQ ID NO 5811; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX polypeptides. The proteins and their associated DNA sequences are used in

XX the treatment, prevention and diagnosis of medical conditions caused by

XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

XX pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

XX P. acnes is also involved in infections of bone, joints and the central

XX nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://wipo.int/pub/published_pct_sequences.

SQ Sequence 322 AA;

Query Match 90.3%; Score 28; DB 22; Length 322;

Best Local Similarity 80.0%; Pred. No. 1.2e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
11:11
Db 148 YTECR 152

RESULT 9

AAG24041 AAG24041 standard; Protein: 435 AA.

AC AAG24041;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27566.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

PM EPI033405-A2.

PD 06-SEP-2000.

PE 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 03-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148568.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 435;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
11:11
Db 34 YTECR 38

RESULT 10
AAG24040
ID AAG24040 standard; Protein: 440 AA.
XX
AC AAG24040;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27565.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PT 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145092.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147320.
PR 06-AUG-1999; 99US-0147330.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152163.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 440;
Best Local Similarity 80.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
11:11
DB 39 YTECR 43

RESULT 11
ABB61563
ID ABB61563 standard; Protein: 441 AA.

AC ABB61563;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 11481.

KW Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL05666.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

PS Disclosure: SEQ ID NO 11481; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 441 AA;

Query Match 90.3%; Score 28; DB 22; Length 441;

Best Local Similarity 80.0%; Pred. No. 1.5e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
11:11
DB 260 YTECR 264

RESULT 12

AAU29351

ID AAU29351 standard; Protein: 476 AA.

XX AAU29351;

XX 18-DEC-2001 (first entry)

DE Novel mar regulated protein (NIMR) #23.

KW mar regulated polypeptide; NIMR; microbial infection; antibacterial.

OS Escherichia coli.

PN WO200170776-A2.

PD 27-SEP-2001.

PF 08-MAR-2001; 2001WO-US07478.

PR 10-MAR-2000; 2000US-188362P.

PA (TUFT) TUFTS COLLEGE.

PI Levy SB, Barbosa TM, Aleksun MN;

DR WPI; 2001-602769/68.

DR N-PSDB; AAS46251.

PT Identifying compounds that modulate a newly identified mar regulated

PT polypeptide activity, useful as antimicrobial compounds, involves

PT contacting the polypeptide with a test compound -

XX Disclosure: Page 302-303; 526pp; English.

XX The invention relates to a method of identifying compounds that modulate

CC a newly identified mar regulated (NIMR) polypeptide activity. The method

CC comprises contacting an NIMR polypeptide with a test compound under

CC interaction conditions, determining the ability of the compound to

CC modulate the activity or expression of the polypeptide, and selecting the

CC modulators. NIMR nucleic acids and polypeptides are used in the treatment

CC of microbial infections, and in screening for modulators of NIMR

CC expression and activity. These modulators can be used to reduce the

CC infectivity of a microbe on a surface, and the virulence of a microbe in

CC a subject suffering from an infection. AAU29329-AAU29379 represent

CC Escherichia coli NIMR amino acid sequences of the invention.

XX Sequence 476 AA;

XX Query Match 90.3%; Score 28; DB 22; Length 476;

XX Best Local Similarity 80.0%; Pred. No. 1.6e+03;

XX Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
11:11
DB 296 YTECR 300

RESULT 13

ABG25044

ID ABG25044 standard; Protein: 493 AA.

XX ABG25044;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #25035.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI WPI: 2001-639362/73.
DR N-PSDB; AAS89231.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
PS Claim 20; SEQ ID NO 55403; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 493 AA;
Query Match 90.3%; Score 28; DB 22; Length 493;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 YTOCR 5
DB 335 YTECR 339
RESULT 14
ABB67974
ID ABB67974 standard; Protein; 652 AA.
XX
XX ABB67974;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 30714.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD

PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI: 2001-656860/75.
DR N-PSDB; ABL12077.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
XX Disclosure; SEQ ID NO 30714; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABU30511), expressed DNA
CC sequences (ABU57737-ABU72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 652 AA;
Query Match 90.3%; Score 28; DB 22; Length 652;
Best Local Similarity 80.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 YTOCR 5
DB 120 YTECR 124
RESULT 15
AAG89761
ID AAG89761 standard; Protein; 856 AA.
XX
XX AAG89761;
AC
XX
XX 26-SEP-2001 (first entry)
DT
XX
XX C glutamicum protein fragment SEQ ID NO: 3515.
DE
XX
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
XX Corynebacterium glutamicum.
OS
XX
XX EPI108790-A2.
PN
XX
XX 20-JUN-2001.
PD
XX
XX 18-DEC-2000; 2000EP-0127688.
PF
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
DR WPI: 2001-376931/40.
DR N-PSDB; AAH64980.
DR

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT

PS Claim 17: SEQ ID NO: 3515; 246pp + Sequence Listing; English.
XX

CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX

SQ Sequence 856 AA;

Query Match

Best Local Similarity 90.3%; Score 28; DB 22; Length 856;

Pred. No. 2.7e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
||:||

Db 126 YTECR 130

Search completed: July 9, 2003, 12:00:07
Job time : 23 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:58:36 ; Search time 11.25 Seconds
(Without alignments)
13.077 Million cell updates/sec

Title: US-09-851-058-3

Perfect score: 31

Sequence: 1 YTCR 5

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/p/tdata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/p/tdata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/p/tdata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/p/tdata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/p/tdata/1/1aa/6C.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	87.1	243	US-09-173-300-22	Sequence 22, Appli
2	27	87.1	452	US-09-035-706-2	Sequence 2, Appli
3	27	87.1	452	US-08-955-841-2	Sequence 2, Appli
4	27	87.1	452	US-09-390-425-2	Sequence 2, Appli
5	27	87.1	452	US-09-566-906-2	Sequence 2, Appli
6	26	83.9	462	US-09-238-796-2	Sequence 2, Appli
7	26	83.9	562	5258502-2	Patent No. 5258502
8	26	83.9	651	US-08-264-101-2	Sequence 2, Appli
9	26	83.9	651	US-08-765-243-2	Sequence 2, Appli
10	26	83.9	651	PCT-US95-07295-2	Sequence 2, Appli
11	26	83.9	734	US-08-765-243-8	Sequence 2, Appli
12	26	83.9	734	PCT-US95-07295-8	Sequence 8, Appli
13	26	83.9	1005	US-08-935-450-2	Sequence 2, Appli
14	26	83.9	1208	US-09-463-702A-2	Sequence 2, Appli
15	26	83.9	1525	US-09-191-647-2	Sequence 2, Appli
16	26	83.9	1525	US-09-540-245A-2	Sequence 2, Appli
17	26	83.9	1525	US-09-540-153-2	Sequence 2, Appli
18	26	83.9	1693	US-08-840-316-1	Sequence 1, Appli
19	26	83.9	1693	US-08-478-507-7	Sequence 1, Appli
20	26	83.9	1693	US-08-809-523-1	Sequence 1, Appli
21	26	83.9	1693	US-09-128-275A-7	Sequence 7, Appli
22	26	83.9	1693	US-08-471-971-1	Sequence 1, Appli
23	26	83.9	1693	US-09-553-427-7	Sequence 7, Appli
24	26	83.9	1693	US-09-402-776-1	Sequence 1, Appli
25	26	83.9	1693	PCT-US93-08849A-1	Sequence 1, Appli
26	26	83.9	1693	PCT-US93-08849-1	Sequence 1, Appli
27	26	83.9	1708	US-09-462-606-2	Sequence 2, Appli

28	26	83.9	4551	US-09-320-878-1	Sequence 1, Appli
29	26	83.9	4613	US-09-105-537-31	Sequence 31, Appli
30	26	83.9	11877	US-09-105-537-6	Sequence 6, Appli
31	25	80.6	20	US-08-484-135-88	Sequence 88, Appli
32	25	80.6	20	US-08-484-635-50	Sequence 50, Appli
33	25	80.6	20	US-08-484-631-50	Sequence 50, Appli
34	25	80.6	102	US-08-827-570-50	Sequence 44, Appli
35	25	80.6	150	US-08-469-260A-44	Sequence 3, Appli
36	25	80.6	155	US-08-867-676-3	Sequence 9, Appli
37	25	80.6	155	US-08-150-203A-9	Sequence 9, Appli
38	25	80.6	155	US-08-454-730-9	Sequence 9, Appli
39	25	80.6	155	US-08-949-788-9	Sequence 9, Appli
40	25	80.6	163	US-08-867-676-1	Sequence 1, Appli
41	25	80.6	222	US-08-867-676-1	Sequence 609, Appli
42	25	80.6	264	US-08-469-260A-76	Sequence 76, Appli
43	25	80.6	295	US-09-799-345-2	Sequence 2, Appli
44	25	80.6	377	US-09-410-464-8	Sequence 8, Appli
45	25	80.6	424	US-08-592-214A-16	Sequence 16, Appli

ALIGNMENTS

```
RESULT 1
US-09-173-300-22
: Sequence 22, Application US/09173300
: Patent No. 6451581
: GENERAL INFORMATION:
: APPLICANT: Falco, Saverio Carl
: APPLICANT: Hiltz, William D.
: APPLICANT: Kinney, Anthony J.
: APPLICANT: Cahoon, Rebecca E.
: APPLICANT: Rafalst, J. Antoni
: TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
: FILE REFERENCE: BB-1126
: CURRENT APPLICATION NUMBER: US/09/173.300
: EARLIER FILING DATE: 1998-10-15
: EARLIER APPLICATION NUMBER: 60/063.423
: NUMBER OF SEQ ID NOS: 54
: SOFTWARE: Microsoft Word Version 7.0A
: SEQ ID NO 22
: LENGTH: 243
: TYPE: PRT
: ORGANISM: Glycine max
: US-09-173-300-22

Query Match      87.1%; Score 27; DB 4; Length 243;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 YTCR 5
Db      68 YTCR 72

RESULT 2
US-09-035-706-2
: Sequence 2, Application US/09035706
: Patent No. 6001622
: GENERAL INFORMATION:
: APPLICANT: Dedhar, Shoukat
: APPLICANT: Hannigan, Greg
: TITLE OF INVENTION: Integrin-Linked Kinase and
: TITLE OF INVENTION: Its Uses
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESS: Bozicevic & Reed, LLP
: STREET: 285 Hamilton Avenue, Suite 200
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94301
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/035,706
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-035-706-2

Query Match
Best Local Similarity 87.1%; Score 27; DB 3; Length 452;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCOR 5
DB 5 FTQCR 9

RESULT 3
US-08-955-841-2
Sequence 2, Application US/08955841
Patent No. 6013782
GENERAL INFORMATION:
APPLICANT: Dedhar, Shoukat
APPLICANT: Hannigan, Greg
TITLE OF INVENTION: Integrin-Linked Kinase and
TITLE OF INVENTION: Its Uses
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,841
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: KIN-2C1P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
```

```
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-841-2

Query Match
Best Local Similarity 87.1%; Score 27; DB 3; Length 452;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCOR 5
DB 5 FTQCR 9

RESULT 4
US-09-390-425-2
Sequence 2, Application US/09390425
Patent No. 6338958
GENERAL INFORMATION:
APPLICANT: Dedhar, Shoukat
APPLICANT: Hannigan, Greg
TITLE OF INVENTION: Integrin-Linked Kinase and Its Uses
FILE REFERENCE: KIN-2CON
CURRENT APPLICATION NUMBER: US/09/390,425
EARLIER FILING DATE: 1999-09-03
EARLIER APPLICATION NUMBER: US60/009,074
EARLIER FILING DATE: 1995-12-21
EARLIER APPLICATION NUMBER: US08/752,345
EARLIER FILING DATE: 1996-11-19
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 452
TYPE: PRT
ORGANISM: H. sapiens
US-09-390-425-2

Query Match
Best Local Similarity 87.1%; Score 27; DB 4; Length 452;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCOR 5
DB 5 FTQCR 9

RESULT 5
US-09-566-906-2
Sequence 2, Application US/09566906
Patent No. 6369205
GENERAL INFORMATION:
APPLICANT: Dedhar, Shoukat
APPLICANT: Hannigan, Greg
TITLE OF INVENTION: Integrin-Linked Kinase and Its Uses
FILE REFERENCE: KIN-2CON
CURRENT APPLICATION NUMBER: US/09/566,906
CURRENT FILING DATE: 2000-05-09
PRIOR APPLICATION NUMBER: 09/390,425
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: US08/752,345
PRIOR FILING DATE: 1996-11-19
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 452
TYPE: PRT
ORGANISM: H. sapiens
```

US-09-566-906-2

Query Match

Best Local Similarity 87.1%; Score 27; DB 4; Length 452;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOCR 5

Db 5 FTQCR 9

RESULT 6

US-09-238-796-2

Sequence 2, Application US/09238796

Patent No. 6074845

GENERAL INFORMATION:

APPLICANT: ALTAR, NAMBI V.

APPLICANT: DISA, JYOTI

TITLE OF INVENTION: BECRLR:BOVINE CALCITONIN RECEPTOR-LIKE

TITLE OF INVENTION: RECEPTOR

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ratner & Prestlia

STREET: P.O. Box 980

CITY: Valley Forge

STATE: PA

COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/238.796

FILING DATE: 28-JAN-1999

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Prestlia, Paul F

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: GP-70599

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-407-0700

TELEFAX: 610-407-0700

TELEX: 846169

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 462 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-238-796-2

Query Match 83.9%; Score 26; DB 3; Length 462;

Best Local Similarity 100.0%; Pred. No. 6.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4

Db 125 YTOC 128

RESULT 7

5258502-2

Patent No. 5258502

APPLICANT: Kuranda, Michael J.

TITLE OF INVENTION: IMMOBILIZATION AND PURIFICATION OF

FUSION PROTEINS USING CHITIN-BINDING ABILITY

NUMBER OF SEQUENCES: 8

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/303,827

FILING DATE: 30-JAN-1989

SEQ ID NO: 2

LENGTH: 562

5258502-2

Query Match

Best Local Similarity 83.9%; Score 26; DB 6; Length 562;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4

Db 549 YTOC 552

RESULT 8

US-08-264-101-2

Sequence 2, Application US/08264101

Patent No. 5693496

GENERAL INFORMATION:

APPLICANT: ALVES, KENNETH

APPLICANT: GUPTA, SUNIL K.

TITLE OF INVENTION: CONTRACEPTIVE VACCINE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: MARY A. APOLLINA

STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE

CITY: RAHWAY

STATE: NJ

COUNTRY: USA

ZIP: 07065

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/264,101

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: APOLLINA, MARY A.

REGISTRATION NUMBER: 34,087

REFERENCE/DOCKET NUMBER: 19244

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908)594-3462

TELEFAX: (908)594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 651 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-264-101-2

Query Match 83.9%; Score 26; DB 1; Length 651;

Best Local Similarity 100.0%; Pred. No. 9.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4

Db 448 YTOC 451

RESULT 9

US-08-765-243-2

Sequence 2, Application US/08765243

Patent No. 5935578

GENERAL INFORMATION:

APPLICANT: ALVES, KENNETH

APPLICANT: GUPTA, SUNIL K.
TITLE OF INVENTION: CONTRACEPTIVE VACCINE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: MARY A. APOLLINA
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
PCT-US95-07295-2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,243
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: APOLLINA, MARY A
REGISTRATION NUMBER: 34,087
REFERENCE/DOCKET NUMBER: 19244Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3462
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-243-2

Query Match 83.9%; Score 26; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4
Db 448 YTOC 451

RESULT 10
PCT-US95-07295-2
GENERAL INFORMATION:
APPLICANT: ALVES, KENNETH
APPLICANT: GUPTA, SUNIL K.
APPLICANT: HOLLIS, GREGORY F.
TITLE OF INVENTION: CONTRACEPTIVE VACCINE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: MARY A. APOLLINA
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07295
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: APOLLINA, MARY A
REGISTRATION NUMBER: 34,087

REFERENCE/DOCKET NUMBER: 19244Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3462
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-07295-2

Query Match 83.9%; Score 26; DB 5; Length 651;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4
Db 448 YTOC 451

RESULT 11
US-08-765-243-8
Sequence 8, Application US/08765243
Patent No. 5935578
GENERAL INFORMATION:
APPLICANT: ALVES, KENNETH
APPLICANT: GUPTA, SUNIL K.
APPLICANT: HOLLIS, GREGORY F.
TITLE OF INVENTION: CONTRACEPTIVE VACCINE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: MARY A. APOLLINA
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,243
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: APOLLINA, MARY A
REGISTRATION NUMBER: 34,087
REFERENCE/DOCKET NUMBER: 19244Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3462
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 734 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-243-8

Query Match 83.9%; Score 26; DB 2; Length 734;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4
Db 531 YTOC 534

RESULT 12

PCT-US95-07295-8

; Sequence 8, Application PC/TUS9507295

; GENERAL INFORMATION:

; APPLICANT: ALVES, KENNETH

; APPLICANT: GUPTA, SUNIL K.

; TITLE OF INVENTION: CONTRACEPTIVE VACCINE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MARY A. APOLLINA

; STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE

; CITY: RAHWAY

; STATE: NJ

; COUNTRY: USA

; ZIP: 07065

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/07295

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: APOLLINA, MARY A

; REGISTRATION NUMBER: 34,087

; REFERENCE/DOCKET NUMBER: 19244Y

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908)594-3462

; TELEFAX: (908)594-4720

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 734 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; PCT-US95-07295-8

Query Match

Best Local Similarity 83.9%; Score 26; DB 5; Length 734;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4

Db 531 YTOC 534

RESULT 13

; Sequence 2, Application US/08935450

; Patent No. 5977311

; GENERAL INFORMATION:

; APPLICANT: Nandabalan, Krishnan

; APPLICANT: Yang, Meijia

; APPLICANT: Schulz, Vincent

; TITLE OF INVENTION: 53BP2 COMPLEXES

; FILE REFERENCE: 7934-054

; CURRENT APPLICATION NUMBER: US/08/935,450

; CURRENT FILING DATE: 1997-09-23

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1005

; TYPE: PRT

; ORGANISM: Homo sapiens

Query Match 83.9%; Score 26; DB 2; Length 1005;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4

Db 917 YTOC 920

RESULT 14

; Sequence 2, Application US/09463702A

; Patent No. 6335435

; GENERAL INFORMATION:

; APPLICANT: AGENE Research Institute, Co., Ltd.

; APPLICANT: HIRAKI AND ASSOCIATES

; APPLICANT: SHIMAMOTO, AKIRO

; APPLICANT: KITAO, SAORI

; APPLICANT: FURUICHI, YASUHIRO

; TITLE OF INVENTION: HUMAN GENE RECQ4 ENCODING HELICASE

; FILE REFERENCE: HIRAI150

; CURRENT APPLICATION NUMBER: US/09/463,702A

; CURRENT FILING DATE: 2000-01-24

; PRIOR APPLICATION NUMBER: PCT/JP98/03114

; PRIOR FILING DATE: 1998-07-10

; PRIOR APPLICATION NUMBER: JAPAN 9/200387

; PRIOR FILING DATE: 1997-07-25

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: Patent In version 3.0

; SEQ ID NO 2

; LENGTH: 1208

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-463-702A-2

Query Match 83.9%; Score 26; DB 4; Length 1208;

Best Local Similarity 80.0%; Pred. No. 1.5e+03;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTOC 5

Db 942 YTOC 946

RESULT 15

; Sequence 2, Application US/09191647

; Patent No. 6046015

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey

; APPLICANT: Brose, Katja

; APPLICANT: Tessier-Lavigne, Marc

; TITLE OF INVENTION: Modulating Robo: Ligand Interactions

; FILE REFERENCE: B98-031-3

; CURRENT APPLICATION NUMBER: US/09/191,647

; CURRENT FILING DATE: 1998-11-13

; EARLIER APPLICATION NUMBER: 60/065,544

; EARLIER FILING DATE: 1997-11-14

; EARLIER APPLICATION NUMBER: 60/081,057

; EARLIER FILING DATE: 1998-04-07

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1525

; TYPE: PRT

; ORGANISM: human

Query Match 83.9%; Score 26; DB 3; Length 1525;

Best Local Similarity 100.0%; Pred. No. 1.9e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4

Db 233 YTOC 236

Thu Jul 10 13:20:58 2003

Search completed: July 9, 2003, 12:03:37
Job time : 12.25 secs

us-09-851-058-3.ra1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:02:11 ; Search time 20.25 Seconds
(Without alignments)
28.746 Million cell updates/sec

Title: US-09-851-058-3
Perfect score: 31
Sequence: 1 YTCOR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues
Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PC1_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/PCT05_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	90.3	818	9 US-10-156-761-9960	Sequence 9960, Ap
2	28	90.3	856	9 US-09-738-626-3515	Sequence 3515, Ap
3	27	87.1	243	12 US-10-027-650-22	Sequence 22, Appl
4	27	87.1	277	9 US-10-106-698-5823	Sequence 5823, Ap
5	27	87.1	307	10 US-09-764-864-1222	Sequence 1222, Ap
6	27	87.1	337	10 US-09-925-300-1708	Sequence 1708, Ap
7	27	87.1	452	10 US-09-840-704-2	Sequence 2, Appl
8	27	87.1	643	9 US-10-198-070-43	Sequence 43, Appl
9	27	87.1	700	9 US-09-906-419-17	Sequence 17, Appl
10	27	87.1	700	9 US-10-119-136-17	Sequence 17, Appl
11	27	87.1	700	9 US-10-119-136-130	Sequence 130, App
12	27	87.1	729	9 US-10-106-698-6364	Sequence 6364, App
13	26	83.9	52	9 US-10-106-698-7543	Sequence 7543, Ap
14	26	83.9	64	10 US-09-864-761-34561	Sequence 34561, A
15	26	83.9	68	9 US-10-083-357-1024	Sequence 1024, Ap
16	26	83.9	139	9 US-10-125-540-521	Sequence 521, App
17	26	83.9	139	10 US-09-764-870-521	Sequence 521, App
18	26	83.9	158	9 US-10-295-403-118	Sequence 118, App
19	26	83.9	342	9 US-10-156-761-8563	Sequence 8563, Ap

20	26	83.9	365	9 US-09-895-913A-264	Sequence 264, App
21	26	83.9	461	9 US-10-051-186-2	Sequence 2, Appl
22	26	83.9	461	9 US-10-225-567A-82	Sequence 82, Appl
23	26	83.9	493	9 US-09-974-298-52	Sequence 52, Appl
24	26	83.9	493	9 US-09-965-528-21	Sequence 21, Appl
25	26	83.9	493	9 US-10-153-668-334	Sequence 334, App
26	26	83.9	560	9 US-10-108-605-315	Sequence 315, App
27	26	83.9	562	10 US-09-801-368-82	Sequence 82, Appl
28	26	83.9	734	9 US-10-054-683-19	Sequence 19, Appl
29	26	83.9	734	9 US-10-205-823-10	Sequence 10, Appl
30	26	83.9	735	9 US-10-205-823-8	Sequence 8, Appl
31	26	83.9	753	9 US-10-205-823-6	Sequence 6, Appl
32	26	83.9	796	9 US-10-028-392-5	Sequence 5, Appl
33	26	83.9	797	9 US-10-270-875-30	Sequence 30, Appl
34	26	83.9	797	9 US-10-270-878-30	Sequence 30, Appl
35	26	83.9	797	9 US-10-270-786-30	Sequence 30, Appl
36	26	83.9	797	9 US-10-270-710-30	Sequence 30, Appl
37	26	83.9	797	9 US-10-270-859-30	Sequence 30, Appl
38	26	83.9	1005	9 US-10-114-091-2	Sequence 2, Appl
39	26	83.9	1529	9 US-09-866-050A-396	Sequence 396, App
40	26	83.9	1529	9 US-10-189-123-67	Sequence 67, Appl
41	26	83.9	1693	9 US-09-851-410-7	Sequence 7, Appl
42	26	83.9	1698	9 US-09-468-147-91	Sequence 91, Appl
43	26	83.9	1708	9 US-09-468-147-166	Sequence 166, App
44	26	83.9	4019	9 US-09-854-133-425	Sequence 425, App
45	26	83.9	4019	9 US-10-144-649A-425	Sequence 425, App

ALIGNMENTS

RESULT 1
US-10-156-761-9960
; Sequence 9960, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9960
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9960
Query Match 90.3%, Score 28; DB 9; Length 818;
Best Local Similarity 80.0%; Pred. No. 8.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CY 1 YTCOR 5
DB 133 YTCOR 137
RESULT 2
US-09-738-626-3515
; Sequence 3515, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIO
APPLICANT: OCHIAI, KEIICHO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 3515
LENGTH: 856
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3515

Query Match 90.3%; Score 28; DB 9; Length 856;
Best Local Similarity 80.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 126 YTECR 130

RESULT 3
US-10-027-450-22
Sequence 22, Application US/10027450
Patent No. US20020102715A1
GENERAL INFORMATION:
APPLICANT: Falco, Saverio Carl
APPLICANT: Hiltz, William D.
APPLICANT: Kinney, Anthony J.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, J. Antoni
TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
FILE REFERENCE: BB-1126
CURRENT APPLICATION NUMBER: US/10/027,450
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/063,423
PRIOR FILING DATE: 1997 October 28
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Microsoft Word Version 7.0A
SEQ ID NO 22
LENGTH: 243
TYPE: PRT
ORGANISM: Glycine max
US-10-027-450-22

Query Match 87.1%; Score 27; DB 12; Length 243;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 68 YSOCR 72

RESULT 4
US-10-106-698-5823
Sequence 5823, Application US/10106698
Patent No. US20030109690A1

GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
FILE REFERENCE: PA00591
CURRENT APPLICATION NUMBER: US/10/106,698
PRIOR FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 5823
LENGTH: 277
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (45)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5823

Query Match 87.1%; Score 27; DB 9; Length 277;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 157 YTKCR 161

RESULT 5
US-09-764-864-1222
Sequence 1222, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
PRIOR FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1222
LENGTH: 307
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (283)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (293)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1222

Query Match 87.1%; Score 27; DB 10; Length 307;
Best Local Similarity 80.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 16 YTOCO 20

RESULT 6
US-09-925-300-1708
Sequence 1708, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:

APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1708
LENGTH: 337
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (112)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (283)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1708

Query Match 87.1%; Score 27; DB 10; Length 337;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
1111:
DB 16 YTCR 20

RESULT 7
US-09-840-704-2
Sequence 2, Application US/09840704
Patent No. US20020122801A1
GENERAL INFORMATION:
APPLICANT: Dechkar, Shoukat
APPLICANT: Hannigan, Greg
TITLE OF INVENTION: Intergrin-Linked Kinase and Its Uses
FILE REFERENCE: KIN-2CON
CURRENT APPLICATION NUMBER: US/09/840,704
CURRENT FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/566,906
PRIOR FILING DATE: 2000-05-09
PRIOR APPLICATION NUMBER: US08/752,345
PRIOR FILING DATE: 1996-11-19
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 452
TYPE: PRT
ORGANISM: H. sapiens
US-09-840-704-2

Query Match 87.1%; Score 27; DB 10; Length 452;
Best Local Similarity 80.0%; Pred. No. 7.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
1111:
DB 5 YTCR 9

RESULT 8
US-10-198-070-43
Sequence 43, Application US/10198070
Publication No. US20030109437A1

GENERAL INFORMATION:
APPLICANT: AVERBACK, PAUL
APPLICANT: GEMMELL, JACK
TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
TITLE OF INVENTION: CELLS
FILE REFERENCE: 59003,000008
CURRENT APPLICATION NUMBER: US/10/198,070
CURRENT FILING DATE: 2002-07-19
PRIOR APPLICATION NUMBER: 60/306,161
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/306,150
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/331,477
PRIOR FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 125
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 643
TYPE: PRT
ORGANISM: Homo sapiens
US-10-198-070-43

Query Match 87.1%; Score 27; DB 9; Length 643;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
1111:
DB 523 YTCR 527

RESULT 9
US-09-906-419-17
Sequence 17, Application US/09906419
Publication No. US20030037357A1
GENERAL INFORMATION:
APPLICANT: Schockley, Jay
APPLICANT: Schnurr, Judy
APPLICANT: Browse, John
TITLE OF INVENTION: Plant Acyl-CoA Synthetases
FILE REFERENCE: DOM-04679
CURRENT APPLICATION NUMBER: US/09/906,419
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 60/220,474
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn version 3.0
SEQ ID NO 17
LENGTH: 700
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-906-419-17

Query Match 87.1%; Score 27; DB 9; Length 700;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
1111:
DB 570 YTCR 574

RESULT 10
US-10-119-136-17
Sequence 17, Application US/10119136
Publication No. US20030097676A1
GENERAL INFORMATION:
APPLICANT: Schockley, Jay
APPLICANT: Schnurr, Judy
APPLICANT: Browse, John
TITLE OF INVENTION: Plant Acyl-CoA Synthetases
FILE REFERENCE: DOM-04695

```
;; CURRENT APPLICATION NUMBER: US/10/119,136
;; CURRENT FILING DATE: 2002-09-04
;; PRIOR APPLICATION NUMBER: 09/906,419
;; PRIOR FILING DATE: 2001-07-16
;; PRIOR APPLICATION NUMBER: 60/220,474
;; PRIOR FILING DATE: 2000-07-21
;; NUMBER OF SEQ ID NOS: 132
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 17
;; LENGTH: 700
;; TYPE: PRT
;; ORGANISM: Arabidopsis thaliana
US-10-119-136-17
```

```
Query Match
Best Local Similarity 87.1%; Score 27; DB 9; Length 700;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 570 YTKCR 574
```

```
RESULT 11
US-10-119-136-130
; Sequence 130, Application US/10119136
; Publication No. US20030097676a1
; GENERAL INFORMATION:
; APPLICANT: Shockey, Jay
; APPLICANT: Schuur, Judy
; APPLICANT: Browne, John
; TITLE OF INVENTION: Plant Acyl-CoA Synthetases
; FILE REFERENCE: DOW-04695
; CURRENT APPLICATION NUMBER: US/10/119,136
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/906,419
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/220,474
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-119-136-130
```

```
Query Match
Best Local Similarity 87.1%; Score 27; DB 9; Length 700;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 570 YTKCR 574
```

```
RESULT 12
US-10-106-698-6364
; Sequence 6364, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
```

```
;; SOFTWARE: PatentIn Ver. 3.0
;; SEQ ID NO 6364
;; LENGTH: 729
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: MISC_FEATURE
;; LOCATION: (30)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: MISC_FEATURE
;; LOCATION: (196)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6364
```

```
Query Match
Best Local Similarity 87.1%; Score 27; DB 9; Length 729;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 274 YTOCR 278
```

```
RESULT 13
US-10-106-698-7543
; Sequence 7543, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7543
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-7543
```

```
Query Match
Best Local Similarity 83.9%; Score 26; DB 9; Length 52;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 29 YTOCR 33
```

```
RESULT 14
US-09-864-761-34561
; Sequence 34561, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
```

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34561
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004689.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 19
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 14
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 16
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 14
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 13
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 15
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 12
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 12
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 14
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 11
OTHER INFORMATION: SWISSPROT HIT: Q53176, EVALU 1.90e+00
US-09-854-761-34561

Query Match 83.9%; Score 26; DB 10; Length 64;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4
1111
Db 52 YTOC 55

RESULT 15

US-10-083-357-1024
Sequence 1024, Application US/10083357
Publication No. US20030054370A1
GENERAL INFORMATION:
APPLICANT: Qiantong Zeng et al.
TITLE OF INVENTION: Systemic Discovery of New Genes
FILE REFERENCE: 032796-090
CURRENT APPLICATION NUMBER: US/10/083,357
CURRENT FILING DATE: 2002-02-27

NUMBER OF SEQ ID NOS: 1346
SEQ ID NO 1024
LENGTH: 68
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1024

Query Match 83.9%; Score 26; DB 9; Length 68;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOC 4
1111
Db 27 YTOC 30

Search completed: July 9, 2003, 12:17:59
Job time : 21.25 secs

This Page Blank (uspto)

Th: * (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:00:11 ; Search time 143.25 Seconds

(without alignments)
22.504 Million cell updates/sec

Title: US-09-851-058-3

Perfect score: 31

Sequence: 1 YTCR 5

Scoring table: BLOSUM62

Searched: 4569144 seqs, 64473110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:

```
1: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US09.COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US10.COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US10.COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US10.COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US10.COMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	39	US-09-513-996A-11153	Sequence 11153, A
2	31	100.0	155	US-09-791-537-101157	Sequence 101157, A
3	31	100.0	155	US-09-791-537-129198	Sequence 129198, A
4	31	100.0	155	US-09-791-537-129201	Sequence 129201, A
5	31	100.0	155	US-09-791-537-129589	Sequence 129589, A
6	31	100.0	157	US-09-791-537-53532	Sequence 53532, A

7	31	100.0	378	19	US-09-570-581A-1947	Sequence 1947, Ap
8	31	100.0	378	19	US-09-573-655A-98	Sequence 98, Appl
9	31	100.0	378	19	US-09-573-655B-98	Sequence 98, Appl
10	31	100.0	555	27	US-60-360-039-12740	Sequence 12740, A
11	31	100.0	676	27	US-60-360-039-3268	Sequence 3268, Ap
12	31	100.0	720	18	US-09-417-507-38363	Sequence 38363, A
13	28	90.3	44	1	PCT-US01-18569-3964	Sequence 3964, Ap
14	28	90.3	44	26	US-10-264-049-3964	Sequence 3964, Ap
15	28	90.3	53	27	US-60-142-896-1135	Sequence 1135, Ap
16	28	90.3	57	27	US-60-361-742-1135	Sequence 1135, Ap
17	28	90.3	105	16	US-09-270-767-60526	Sequence 60526, A
18	28	90.3	105	16	US-09-270-849B-190319	Sequence 190319, A
19	28	90.3	129	18	US-09-489-039A-10711	Sequence 10711, A
20	28	90.3	148	21	US-09-791-537-103013	Sequence 103013, A
21	28	90.3	177	26	US-10-221-279-9040	Sequence 9040, Ap
22	28	90.3	291	20	US-09-604-603A-114	Sequence 114, App
23	28	90.3	291	27	US-60-167-217-20509	Sequence 20509, A
24	28	90.3	291	27	US-60-173-464-16715	Sequence 16715, A
25	28	90.3	321	16	US-09-270-767-45035	Sequence 45035, A
26	28	90.3	336	21	US-09-791-537-36108	Sequence 36108, A
27	28	90.3	401	27	US-60-360-039-4107	Sequence 4107, Ap
28	28	90.3	435	19	US-09-513-996A-27566	Sequence 27566, A
29	28	90.3	440	19	US-09-513-996A-27565	Sequence 27565, A
30	28	90.3	441	20	US-09-614-150-11481	Sequence 11481, A
31	28	90.3	441	21	US-09-791-537-17513	Sequence 17513, A
32	28	90.3	441	27	US-60-191-637-11513	Sequence 11513, A
33	28	90.3	441	27	US-60-191-681-9005	Sequence 9005, Ap
34	28	90.3	471	21	US-09-791-537-10483	Sequence 10483, A
35	28	90.3	471	21	US-09-791-537-14387	Sequence 14387, A
36	28	90.3	476	21	US-09-791-537-143801	Sequence 143801, A
37	37	90.3	476	22	US-09-801-563-43	Sequence 43, Appl
38	28	90.3	493	1	PCT-US01-08631-55403	Sequence 55403, A
39	28	90.3	498	27	US-60-360-039-3846	Sequence 3846, Ap
40	28	90.3	652	27	US-09-614-150-30714	Sequence 30714, A
41	28	90.3	652	27	US-60-191-637-30271	Sequence 30271, A
42	28	90.3	667	23	US-09-902-540-15678	Sequence 15678, A
43	28	90.3	780	27	US-60-147-189-665	Sequence 665, App
44	28	90.3	818	21	US-09-791-537-110137	Sequence 110137, A
45	28	90.3	840	21	US-09-791-537-7312	Sequence 7312, Ap

ALIGNMENTS

```
RESULT 1
US-09-513-996A-11153
; Sequence 11153, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513, 996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 11153
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: UNSURE
; LOCATION: 1..39 /
; OTHER INFORMATION: any n or Xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..39 / Ceres Seq. ID 1035338
US-09-513-996A-11153
```

Query Match 100.0%; Score 31; DB 19; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 YTCR 5
|||||

DB 14 YTOCR 18

RESULT 2

US-09-791-537-101157
; Sequence 101157, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 101157
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-101157

Query Match 100.0%; Score 31; DB 21; Length 155;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5

DB 115 YTOCR 119

RESULT 3

US-09-791-537-129198
; Sequence 129198, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129198
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-129198

Query Match 100.0%; Score 31; DB 21; Length 155;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5

DB 115 YTOCR 119

RESULT 4

US-09-791-537-129201
; Sequence 129201, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129201
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-129201

Query Match 100.0%; Score 31; DB 21; Length 155;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5

DB 115 YTOCR 119

RESULT 5

US-09-791-537-129589
; Sequence 129589, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129589
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-129589

Query Match 100.0%; Score 31; DB 21; Length 155;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5

DB 115 YTOCR 119

RESULT 6

US-09-791-537-53532
; Sequence 53532, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53532
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Cricetus griseus
US-09-791-537-53532

Query Match 100.0%; Score 31; DB 21; Length 157;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5

DB 117 YTOCR 121


```
RESULT 7
US-09-570-581A-1947
; Sequence 1947, Application US/09570581A
; GENERAL INFORMATION:
; APPLICANT: Ceres Inc.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0875P
; CURRENT APPLICATION NUMBER: US/09/570,581A
; CURRENT FILING DATE: 2001-05-12
; NUMBER OF SEQ ID NOS: 1950
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1947
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-570-581A-1947

Query Match
Best Local Similarity 100.0%; Score 31; DB 19; Length 378;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
    |||||
Db 194 YTCR 198

RESULT 8
US-09-573-655A-98
; Sequence 98, Application US/09573655A
; GENERAL INFORMATION:
; APPLICANT: SOLOVYEV, Victor and TROUKHAN, Maxim
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0876P
; CURRENT APPLICATION NUMBER: US/09/573,655A
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 3280
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 98
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-573-655A-98

Query Match
Best Local Similarity 100.0%; Score 31; DB 19; Length 378;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
    |||||
Db 194 YTCR 198

RESULT 9
US-09-573-655B-98
; Sequence 98, Application US/09573655B
; GENERAL INFORMATION:
; APPLICANT: SOLOVYEV, Victor and TROUKHAN, Maxim
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0876P
; CURRENT APPLICATION NUMBER: US/09/573,655B
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 3281
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 98
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-573-655B-98

Query Match
Best Local Similarity 100.0%; Score 31; DB 19; Length 378;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
    |||||
Db 194 YTCR 198

RESULT 10
US-60-360-039-12740
; Sequence 12740, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12740
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-60-360-039-12740

Query Match
Best Local Similarity 100.0%; Score 31; DB 27; Length 555;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
    |||||
Db 282 YTCR 286

RESULT 11
US-60-360-039-3268
; Sequence 3268, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3268
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(676)
; OTHER INFORMATION: unsure at all Xaa locations
US-60-360-039-3268

Query Match
Best Local Similarity 100.0%; Score 31; DB 27; Length 676;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
    |||||
Db 300 YTCR 304
```

```
RESULT 12
US-09-417-507-38363
; Sequence 38363, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PAT99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; NUMBER OF SEQ ID NOS: 1999-10-14
; SEQ ID NO 38363
; LENGTH: 720
; TYPE: PRT
; ORGANISM: A.fumigatus
US-09-417-507-38363

Query Match
Best Local Similarity 100.0%; Score 31; DB 18; Length 720;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
DB 246 YTECR 250

RESULT 13
PCT-US01-18569-3964
; Sequence 3964, Application PC/TUS0118569
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA133PCT
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,467
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3964
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-18569-3964

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 44;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
DB 25 YTECR 29

RESULT 14
US-10-264-049-3964
; Sequence 3964, Application US/10264049
; GENERAL INFORMATION:
; APPLICANT: Btse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
```

```
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3964
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3964

Query Match
Best Local Similarity 90.3%; Score 28; DB 26; Length 44;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
DB 25 YTECR 29

RESULT 15
US-60-142-896-1135
; Sequence 1135, Application US/60142896
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: CL000052
; CURRENT APPLICATION NUMBER: US/60/142,896
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 1584
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1135
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Drosophila
US-60-142-896-1135

Query Match
Best Local Similarity 80.0%; Score 26; DB 27; Length 53;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
DB 32 YTECR 36

Search completed: July 9, 2003, 12:13:18
Job time : 144.25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 12:00:51 ; Search time 47.25 Seconds
(Without alignments)
27.740 Million cell updates/sec

Title: US-09-851-058-3
Perfect score: 31
Sequence: 1 YTCOR 5

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1221054 seqs, 262139423 residues

Total number of hits satisfying chosen parameters: 1221054

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US12_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US13_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US14_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US15_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US16_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US17_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	5	2	PCT-US02-14369-3
2	31	100.0	5	2	US-09-851-058-3
3	31	100.0	155	12	US-10-219-0518-12561
4	31	100.0	155	12	US-10-219-0518-14520
5	31	100.0	155	12	US-10-219-0518-14520
6	31	100.0	155	12	US-10-219-0518-14520
7	31	100.0	155	12	US-10-219-0518-14520
8	31	100.0	155	12	US-10-219-0518-14520
9	31	100.0	155	12	US-10-219-0518-14520
10	31	100.0	155	12	US-10-219-0518-14520
11	31	100.0	155	12	US-10-219-0518-14520
12	31	100.0	155	12	US-10-219-0518-14520
13	31	100.0	155	12	US-10-219-0518-14520
14	31	100.0	155	12	US-10-219-0518-14520
15	31	100.0	155	12	US-10-219-0518-14520
16	31	100.0	155	12	US-10-219-0518-14520
17	31	100.0	155	12	US-10-219-0518-14520
18	31	100.0	155	12	US-10-219-0518-14520
19	31	100.0	155	12	US-10-219-0518-14520

20	28	90.3	513	2	PCT-US02-32727-29554	Sequence 29554, A
21	28	90.3	513	10	US-09-978-825-29554	Sequence 29554, A
22	28	90.3	818	12	US-10-156-761-9960	Sequence 9960, Ap
23	28	90.3	840	12	US-10-369-493-22262	Sequence 22262, A
24	28	90.3	856	12	US-10-282-122A-53654	Sequence 53654, A
25	27	87.1	5	2	PCT-US02-14369-3	Sequence 2, Appl1
26	27	87.1	5	2	US-09-851-058-2	Sequence 139, App
27	27	87.1	9	12	US-10-394-980-139	Sequence 365, App
28	27	87.1	9	12	US-10-394-980-365	Sequence 27, Appl
29	27	87.1	40	2	PCT-US03-08715-27	Sequence 165002,
30	27	87.1	41	12	US-10-437-963-165002	Sequence 239546,
31	27	87.1	53	12	US-10-424-599-239546	Sequence 145972,
32	27	87.1	60	12	US-10-424-599-145972	Sequence 227392,
33	27	87.1	61	12	US-10-282-122A-52390	Sequence 58955, A
34	27	87.1	61	12	US-10-282-122A-58955	Sequence 1388, Ap
35	27	87.1	61	12	US-10-421-684-1388	Sequence 140723,
36	27	87.1	67	12	US-10-437-963-140723	Sequence 227392,
37	27	87.1	72	12	US-10-424-599-227392	Sequence 2772, Ap
38	27	87.1	87	12	US-10-219-382-2772	Sequence 239127,
39	27	87.1	87	12	US-10-219-382A-2772	Sequence 60, Appl
40	27	87.1	103	12	US-10-236-177-60	Sequence 229646,
41	27	87.1	107	12	US-10-424-599-229646	Sequence 63170, A
42	27	87.1	108	12	US-10-424-599-63170	Sequence 63170, A
43	27	87.1	115	10	US-09-724-676A-63170	Sequence 254138,
44	27	87.1	115	10	US-09-724-676A-63170	
45	27	87.1	117	12	US-10-424-599-254138	

ALIGNMENTS

RESULT 1
PCT-US02-14369-3
Sequence 3, Application PC/TUS0214369

GENERAL INFORMATION:
APPLICANT: Parker, Kenneth
APPLICANT: Nadler, Timothy
APPLICANT: Vella, George
APPLICANT: Huang, Yulin
APPLICANT: Abersold, Rudolf
APPLICANT: Smolka, Marcus
TITLE OF INVENTION: Process for Analyzing Protein Samples
FILE REFERENCE: SYP-172
CURRENT APPLICATION NUMBER: PCT/US02/14369
CURRENT FILING DATE: 2003-02-05
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent version 3.1
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptides
PCT-US02-14369-3

Query Match 100.0% Score 31; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.1e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCOR 5
DB 1 YTCOR 5

RESULT 2
US-09-851-058-3
Sequence 3, Application US/09851058
GENERAL INFORMATION:
APPLICANT: Parker, Kenneth
APPLICANT: Nadler, Timothy
APPLICANT: Vella, George
APPLICANT: Huang, Yulin
APPLICANT: Abersold, Rudolf

```
APPLICANT: Smolka, Marcus
TITLE OF INVENTION: Process for Analyzing Protein Samples
FILE REFERENCE: SYP-172
CURRENT APPLICATION NUMBER: US/09/851,058
CURRENT FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptides
US-09-851-058-3
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 10; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 YTCR 5
Db 1 YTCR 5
```

```
RESULT 3
US-10-219-051B-12561
Sequence 12561, Application US/10219051B
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation doing business as Massachusetts General
APPLICANT: Hospital / Bayer AG
TITLE OF INVENTION: Nucleotide sequences involved in pain
FILE REFERENCE: Lea 35693 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/219,051B
CURRENT FILING DATE: 2003-05-09
PRIOR APPLICATION NUMBER: US 60/312,147
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: US 60/346,382
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/333,347
PRIOR FILING DATE: 2001-11-26
NUMBER OF SEQ ID NOS: 14715
SOFTWARE: Perl script
SEQ ID NO 12561
LENGTH: 155
TYPE: PRT
ORGANISM: Rattus norvegicus
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: SWISS-Prot / P70709
DATABASE ENTRY DATE: 1998-07-15
US-10-219-051B-12561
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 12; Length 155;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 YTCR 5
Db 115 YTCR 119
```

```
RESULT 4
US-10-219-051B-14320
Sequence 14320, Application US/10219051B
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation doing business as Massachusetts General
APPLICANT: Hospital / Bayer AG
TITLE OF INVENTION: Nucleotide sequences involved in pain
FILE REFERENCE: Lea 35693 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/219,051B
CURRENT FILING DATE: 2003-05-09
PRIOR APPLICATION NUMBER: US 60/312,147
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: US 60/346,382
```

```
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/333,347
PRIOR FILING DATE: 2001-11-26
NUMBER OF SEQ ID NOS: 14715
SOFTWARE: Perl script
SEQ ID NO 14320
LENGTH: 155
TYPE: PRT
ORGANISM: Rattus norvegicus
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: SWISS-Prot / P70709
DATABASE ENTRY DATE: 1998-07-15
US-10-219-051B-14320
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 12; Length 155;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 YTCR 5
Db 115 YTCR 119
```

```
RESULT 5
US-10-437-963-157136
Sequence 157136, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Plant Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 157136
LENGTH: 372
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_56737C.1.pcp
US-10-437-963-157136
```

```
Query Match
Best Local Similarity 100.0%; Score 31; DB 12; Length 372;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 YTCR 5
Db 206 YTCR 210
```

```
RESULT 6
US-10-369-493-12740
Sequence 12740, Application US/10369493
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
```

;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 12740
;; LENGTH: 555
;; TYPE: PRT
;; ORGANISM: Aspergillus nidulans
US-10-369-493-12740

Query Match
Best Local Similarity 100.0%; Score 31; DB 12; Length 555;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOCR 5
|||||
DB 282 YTOCR 286

RESULT 7
US-10-369-493-3268
;; Sequence 3268, Application US/10369493
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Chen, Xianfeng
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; CURRENT FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 3268
;; LENGTH: 676
;; TYPE: PRT
;; ORGANISM: Neurospora crassa
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(676)
;; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3268

Query Match
Best Local Similarity 100.0%; Score 31; DB 12; Length 676;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOCR 5
|||||
DB 300 YTOCR 304

RESULT 8
US-10-446-203-10711
;; Sequence 10711, Application US/10446203
;; GENERAL INFORMATION:
;; APPLICANT: Gary Breton et. al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;; FILE REFERENCE: 2709.2004001
;; CURRENT APPLICATION NUMBER: US/10/446,203
;; CURRENT FILING DATE: 2003-05-27
;; PRIOR APPLICATION NUMBER: US/09/489,039
;; PRIOR FILING DATE: 2000-01-27
;; PRIOR APPLICATION NUMBER: US 60/117,747
;; PRIOR FILING DATE: 1999-01-29
;; NUMBER OF SEQ ID NOS: 14342
;; SEQ ID NO 10711
;; LENGTH: 129
;; TYPE: PRT
;; ORGANISM: Klebsiella pneumoniae
US-10-446-203-10711

Query Match
Best Local Similarity 90.3%; Score 28; DB 12; Length 129;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOCR 5
|||||
DB 108 YTOCR 112

RESULT 9
US-10-282-122A-73614
;; Sequence 73614, Application US/10282122A
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Liangsu
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Malone, Cheryl
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl
;; APPLICANT: Zyskind, Judith
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John
;; APPLICANT: Carr, Grant
;; APPLICANT: Yamamoto, Robert
;; APPLICANT: Forsyth, R.
;; APPLICANT: Xu, H
;; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;; FILE REFERENCE: ELITRA.034A
;; CURRENT APPLICATION NUMBER: US/10/282,122A
;; CURRENT FILING DATE: 2003-02-20
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 73614
;; LENGTH: 304
;; TYPE: PRT
;; ORGANISM: Salmonella paratyphi A
US-10-282-122A-73614

Query Match
Best Local Similarity 90.3%; Score 28; DB 12; Length 304;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTOCR 5
|||||
DB 167 YTOCR 171

RESULT 10
US-10-282-122A-75073
;; Sequence 75073, Application US/10282122A
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Liangsu

```

: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl
: APPLICANT: Zyskind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: EUTRA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: PRIOR FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: Remaining prior Application data removed - See file wrapper or PALM.
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 75073
: LENGTH: 304
: TYPE: PRT
: ORGANISM: Salmonella typhimurium
: US-10-282-122A-75073

Query Match
Best Local Similarity 90.3%; Score 28; DB 12; Length 304;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
Db 167 YTCR 171

RESULT 11
US-10-282-122A-75805
: Sequence 75805, Application US/10282122A
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl
: APPLICANT: Zyskind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: EUTRA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: CURRENT FILING DATE: 2003-02-20
```

```

: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: Remaining prior Application data removed - See file wrapper or PALM.
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 75805
: LENGTH: 304
: TYPE: PRT
: ORGANISM: Salmonella typhl
: US-10-282-122A-75805

Query Match
Best Local Similarity 90.3%; Score 28; DB 12; Length 304;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
Db 167 YTCR 171

RESULT 12
US-10-437-963-177062
: Sequence 177062, Application US/10437963
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Wu, Wei
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Barbazuk, Brad
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
: FILE REFERENCE: 38-21(5321)B
: CURRENT APPLICATION NUMBER: US/10/437,963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 177062
: LENGTH: 313
: TYPE: PRT
: ORGANISM: Oryza sativa
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT4530_74751C.1.pcp
: US-10-437-963-177062

Query Match
Best Local Similarity 90.3%; Score 28; DB 12; Length 313;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
Db 196 YTCR 200
```

```

RESULT 13
US-10-282-122A-59088
: Sequence 59088, Application US/10282122A
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl
: APPLICANT: Zyskind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: APPLICANT: Xu, H.
: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: ELITRA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: CURRENT FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: Remaining Prior Application data removed - See file wrapper or PALM.
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 59088
: LENGTH: 315
: TYPE: PRT
: ORGANISM: Klebsiella pneumoniae
: US-10-282-122A-59088

Query Match          90.3%; Score 28; DB 12; Length 315;
Best Local Similarity 80.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 YTOCR 5
      |||:|
      167 YTOCR 171

DB      167 YTOCR 171

RESULT 14
PCT-US02-32727-5811
: Sequence 5811, Application PC/TUS0232727
: GENERAL INFORMATION:
: APPLICANT: Mitcham, Jennifer
: APPLICANT: Skeiky, Yasir
: APPLICANT: Persing, David
: APPLICANT: Bhatia, Ajay
: APPLICANT: Maisonneuve, Jean Francois
: APPLICANT: Zhang, Yanni
: APPLICANT: Wang, Sigling
: APPLICANT: Jen, Shyian
: APPLICANT: Lodes, Michael
: APPLICANT: Benson, Darin
: APPLICANT: Jones, Robert

```

```

: APPLICANT: Carter, Darick
: APPLICANT: Barth, Brenda
: APPLICANT: Douglass, John
: TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
: FILE REFERENCE: 210121.514C1
: CURRENT APPLICATION NUMBER: PCT/US02/32727
: CURRENT FILING DATE: 2002-10-11
: NUMBER OF SEQ ID NOS: 30992
: SEQ ID NO 5811
: LENGTH: 322
: TYPE: PRT
: ORGANISM: Propionl acnes
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (262)
: OTHER INFORMATION: Xaa - Any Amino Acid
: PCT-US02-32727-5811

Query Match          90.3%; Score 28; DB 2; Length 322;
Best Local Similarity 80.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 YTOCR 5
      |||:|
      148 YTECR 152

DB      148 YTECR 152

RESULT 15
US-09-978-825-5811
: Sequence 5811, Application US/09978825
: GENERAL INFORMATION:
: APPLICANT: Mitcham, Jennifer
: APPLICANT: Skeiky, Yasir
: APPLICANT: Persing, David
: APPLICANT: Bhatia, Ajay
: APPLICANT: Maisonneuve, Jean Francois
: APPLICANT: Zhang, Yanni
: APPLICANT: Wang, Sigling
: APPLICANT: Jen, Shyian
: APPLICANT: Lodes, Michael
: APPLICANT: Benson, Darin
: APPLICANT: Jones, Robert
: APPLICANT: Carter, Darick
: APPLICANT: Barth, Brenda
: APPLICANT: Douglass, John
: TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
: FILE REFERENCE: 210121.514C1
: CURRENT APPLICATION NUMBER: US/09/978,825
: CURRENT FILING DATE: 2003-01-29
: NUMBER OF SEQ ID NOS: 30992
: SEQ ID NO 5811
: LENGTH: 322
: TYPE: PRT
: ORGANISM: Propionl acnes
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (262)
: OTHER INFORMATION: Xaa - Any Amino Acid
: US-09-978-825-5811

Query Match          90.3%; Score 28; DB 10; Length 322;
Best Local Similarity 80.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 YTOCR 5
      |||:|
      148 YTECR 152

DB      148 YTECR 152

Search completed: July 9, 2003, 12:16:33
Job time : 48.25 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:58:06 ; Search time 9.25 Seconds
(Without alignments)
51.965 Million cell updates/sec

Title: US-09-851-058-3

Perfect score: 31
Sequence: 1 YTCGR 5

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	155	2 JC6159	eosinophil-associa
2	31	100.0	156	2 JC6150	eosinophil-associa
3	31	100.0	378	2 B85059	hypothetical prote
4	28	90.3	304	2 AF0730	probable membrane
5	28	90.3	336	2 F81348	ABC transport syst
6	28	90.3	432	2 D64512	hypothetical prote
7	28	90.3	472	2 B82492	tryptophanase vca0
8	28	90.3	476	1 W2EC	tryptophanase (EC
9	28	90.3	476	2 E91209	tryptophanase [imp
10	28	90.3	476	2 H86055	tryptophanase [imp
11	28	90.3	818	2 T35872	DNA gyrase-like pr
12	28	90.3	840	2 A27832	cell division cont
13	28	90.3	872	2 AB1914	DNA gyrase A chain
14	28	90.3	902	2 AF3362	DNA topoisomerase
15	28	90.3	906	2 AB2762	DNA gyrase subunit
16	28	90.3	919	2 C87445	DNA gyrase subunit
17	28	90.3	930	2 A97543	(Y09073) DNA gyras
18	27	87.1	61	2 B64683	ribosomal protein
19	27	87.1	61	2 B71834	ribosomal protein
20	27	87.1	89	2 AB4867	probable trypsin i
21	27	87.1	99	2 S65651	trypsin inhibitor
22	27	87.1	113	2 C24773	protein-tyrosine k
23	27	87.1	221	2 T20781	hypothetical prote
24	27	87.1	262	2 T33408	hypothetical prote
25	27	87.1	262	2 T36968	protein H34124.2 [
26	27	87.1	331	2 T36968	conserved phytoene
27	27	87.1	341	2 B82790	serine/threonine-s
28	27	87.1	451	1 S68455	probable nucleosid
29	27	87.1	516	2 G84442	

30	27	87.1	729	2 T23474	hypothetical prote
31	27	87.1	809	2 A12747	conserved hypotet
32	27	87.1	873	2 H97528	hypothetical prote
33	27	87.1	947	2 T01238	hypothetical prote
34	27	87.1	956	2 AB9153	protein C24B5.3 [1
35	27	87.1	1043	2 T19734	hypothetical prote
36	26	83.9	64	2 E82914	ribosomal protein
37	26	83.9	78	2 T17891	hypothetical prote
38	26	83.9	100	2 S55163	hypothetical prote
39	26	83.9	105	2 E32716	ig heavy chain C r
40	26	83.9	132	2 T20463	hypothetical prote
41	26	83.9	145	2 A64010	hypothetical prote
42	26	83.9	159	2 T51840	RING-H2 finger pro
43	26	83.9	232	2 F70488	hypothetical prote
44	26	83.9	233	2 J00097	hypothetical 26K p
45	26	83.9	236	2 T52494	hypothetical prote

ALIGNMENTS

RESULT 1
JC6159
eosinophil-associated ribonuclease (EC 3.1.1.1) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 22-Jun-1999
C:Accession: JC6159; PC6041
R:Larson, K.A.; Olson, E.V.; Madden, B.J.; Gielch, G.J.; Lee, N.A.; Lee, J.J.
Proc. Natl. Acad. Sci. U.S.A. 93, 12370-12375, 1996
A:Title: Two highly homologous ribonuclease genes expressed in mouse eosinophils
A:Reference number: JC6159; MUID:97057249; PMID:8901588
A:Accession: JC6159
A:Molecule type: mRNA
A:Residues: 1-155 <LAR1>
A:Cross-references: GB:U72032; NID:g1695898; PIDN:AAB37786.1; PID:g1695899
A:Accession: PC6041
A:Molecule type: protein
A:Residues: 52-75;137-149 <LAR2>
A:Experimental source: bone marrow
C:Genetics:
A:Gene: mear-1
C:Superfamily: pancreatic ribonuclease
C:Keywords: bone; disulfide bond; glycoprotein; hydrolase; RNA binding
F:1-77/Domain: signal sequence #status predicted <SIG>
F:28-155/Product: eosinophil-associated ribonuclease 1 #status predicted <MAT>
F:38,62,150/Active site: His, Lys, His #status predicted
F:41,96,107,114/Binding site: carboxylate (Asn) (covalent) #status predicted
F:47-106,61-118,79-133,86-94/Disulfide bonds: #status predicted

Query Match 100.0%; Score 31; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. NO. 19;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCGR 5
DB 115 YTCGR 119

RESULT 2
JC6160
eosinophil-associated ribonuclease (EC 3.1.1.1) 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 22-Jun-1999
C:Accession: JC6160; PC6042
R:Larson, K.A.; Olson, E.V.; Madden, B.J.; Gielch, G.J.; Lee, N.A.; Lee, J.J.
Proc. Natl. Acad. Sci. U.S.A. 93, 12370-12375, 1996
A:Title: Two highly homologous ribonuclease genes expressed in mouse eosinophils
A:Reference number: JC6159; MUID:97057249; PMID:8901588
A:Accession: JC6160
A:Molecule type: mRNA
A:Residues: 1-156 <LAR1>
A:Cross-references: GB:U72031; NID:g1695896; PIDN:AAB37785.1; PID:g1695897
A:Experimental source: bone marrow

A:Accession: PC6042
A:Molecule type: protein
A:Residues: 52-77;138-150 <LAR2>
C:Genetics:

A:Gene: near-2

C:Superfamily: pancreatic ribonuclease

C:Keywords: bone; disulfide bond; glycoprotein; hydrolase; RNA binding

F:1-27/Domain: signal sequence #status predicted <STG>

F:28-156/Product: eosinophil-associated ribonuclease 2 #status predicted <MAT>

F:38,62,151/Active site: His, Lys, His #status predicted

F:47-106,61-119,79-134,86-94/Disulfide bonds: #status predicted

F:89,96,107/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 156;
Pred. No. 19;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
|||||
DB 116 YTCR 120

RESULT 3

B85059
hypothetical protein AT4g04690 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C:Accession: B85059

R:Annotation: The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: B85059

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-378 <STO>

A:Cross-references: GB:NC_001268; NID:g7267227; PIDN:CAB80834.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g04690

A:Map position: 4

Query Match

Best Local Similarity 100.0%; Score 31; DB 2; Length 378;
Pred. No. 39;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
|||||
DB 194 YTCR 198

RESULT 4

AF0730
Probable membrane protein STY2000 [imported] - Salmonella enterica subsp. enterica serov

C:Species: Salmonella enterica subsp. enterica serovar typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001

C:Accession: AF0730

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,

S.; Moutle, S.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AF0730

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-304 <PAR>

C:Genetics:

A:Cross-references: GB:AL513382; PIDN:CA005544.1; PID:g16503045; GSPDB:GN00176

A:Gene: STY2000

Query Match

Best Local Similarity 90.3%; Score 28; DB 2; Length 304;

Best Local Similarity 80.0%; Pred. No. 1,4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
|||||
DB 167 YTCR 171

RESULT 5

F81348
ABC transport system ATP-binding protein Cj0774c [imported] - Campylobacter jejuni (s

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: F81348

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kellley, J.M.; Churcher, C.; Basham, D.; Chl

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Bar

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: AB1250; MUID:20150912; PMID:10688204

A:Accession: F81348

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-336 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:g9968128; PIDN:CAB73039.1; PID:g9696

C:Genetics:

A:Superfamily: probable ABC-type transport protein abc; ATP-binding cassette homology

Query Match

Best Local Similarity 90.3%; Score 28; DB 2; Length 336;
Pred. No. 1,5e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
|||||
DB 110 YTCR 114

RESULT 6

D64512
hypothetical protein MJECL21 - Methanococcus jannaschii plasmid pURB800

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 04-Mar-2000

C:Accession: D64512

R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: D64512

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-432 <BDL>

A:Cross-references: GB:L77118; NID:g1500644; TIGR:MJECL21; PIDN:AC37092.1; PID:g1522

C:Genetics:

A:Map position: ECLREV24593-23295

A:Gene: plasmid

A:Note: this stable 58-kilobase pair plasmid is also designated ECL (large extrachrom

C:Superfamily: Methanococcus jannaschii plasmid pURB800 hypothetical protein MJECL21

Query Match

Best Local Similarity 90.3%; Score 28; DB 2; Length 432;
Pred. No. 1,8e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5
|||||
DB 299 YTCR 303

RESULT 7

B82492

tryptophanase VCA0161 [Imported] - Vibrio cholerae (strain M16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82492
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
I., R.R.; Melakalos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82492
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <HE>
A:Cross-references: GB:AE004357; GB:AE003853; NID:99657547; PIDN:AAF96074.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain M16961; biotype El Tor
C:Genetics:
A:Gene: VCA0161
A:Map position: 2
C:Superfamily: tryptophanase

Query Match 90.3%; Score 28; DB 2; Length 472;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
Db 291 YTECR 295

RESULT 8

tryptophanase (EC 4.1.99.1) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 14-Nov-1993 #sequence_revision 05-Dec-1997 #text_change 01-Mar-2002
C:Accession: E55173; A91789; I41097; I69358; A01136
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
A.; Rose, D.J.; Mau, B.; Shaio, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E55173
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-476 <BAT>
A:Cross-references: GB:AE000448; GB:U00096; NID:91790142; PIDN:AAC76731.1; PID:91790144;
A:Experimental source: strain K-12, substrain MG1655
R:Deeley, M.C.; Yanofsky, C.
J. Bacteriol. 147, 787-796, 1981
A:Title: Nucleotide sequence of the structural gene for tryptophanase of Escherichia coli
A:Reference number: A91789; MUID:82007678; PMID:6268608
A:Accession: A91789
A:Molecule type: DNA
A:Residues: 6-141, 'T', 143, 'OG', 146-383, 'TG', 386-476 <DEE>
A:Experimental source: strain K-12
A>Note: the authors translated the codons GAT, ACG, and CAG for residues 142, 144, and 1
R:Kagamiyama, H.; Matsubara, H.; Spell, E.E.
J. Biol. Chem. 247, 1576-1586, 1972
A:Title: The chemical structure of tryptophanase from Escherichia coli. III. Isolation &
A:Reference number: A92100; MUID:72134434; PMID:4551944
A:Contents: annotation; sequences of tryptic peptides; strain K-12
A>Note: Lys-275 binds pyridoxal 5'-phosphate
R:Tokushige, M.; Tsujimoto, N.; Oda, T.; Honda, T.; Yumoto, N.; Ito, S.; Yamamoto, M.; K
Biochimie 71, 711-720, 1989
A:Title: Role of cysteine residues in tryptophanase for monovalent cation-induced active
A:Reference number: I41097; MUID:89323226; PMID:2502187
A:Accession: I41097
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 6-383, 'TG', 386-476 <RES>
A:Cross-references: EMBL:X15974; NID:941935; PIDN:CAA34096.1; PID:941936
R:Stewart, V.J.; Yanofsky, C.
J. Bacteriol. 164, 731-740, 1985

A:Title: Evidence for transcription antitermination control of tryptophanase operon e
A:Reference number: I54862; MUID:86033634; PMID:3902796
A:Accession: I69358
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 6-27 <RE2>
A:Cross-references: GB:M11990; NID:9147999; PIDN:AAA24679.1; PID:9148001
C:Genetics:
A:Gene: tnaA
A:Map position: 83 min
C:Complex: homotetramer
C:Function:
A:Description: catalyzes the degradation of tryptophan to indole, pyruvate, and ammon
C:Superfamily: tryptophanase
C:Keywords: carbon-carbon lyase; homotetramer

Query Match 90.3%; Score 28; DB 1; Length 476;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
Db 296 YTECR 300

RESULT 9

tryptophanase [Imported] - Escherichia coli (strain O157:H7, substrain RMD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: E91209
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shibb, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E91209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <HAT>
A:Cross-references: GB:BA000007; PIDN:BA838068.1; PID:913364120; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs4645
C:Superfamily: tryptophanase

Query Match 90.3%; Score 28; DB 2; Length 476;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
Db 296 YTECR 300

RESULT 10

tryptophanase [Imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H86055
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Hiller, L.; Grobeck, E.J.; Davis, N.W.; Llm, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H86055
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <STO>
A:Cross-references: GB:AE005174; NID:912518553; PIDN:AAG58908.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:

A:Gene: tnaA
C:Superfamily: tryptophanase

Query Match 90.3%; Score 28; DB 2; Length 476;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 296 YTECR 300

RESULT 11

T35872

DNA gyrase-like protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000

C:Accession: T35872

R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Randal, M.A.

submitted to the EMBL Data Library, November 1997

A:Reference number: Z21592

A:Accession: T35872

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-818 <OLIG>

A:Cross-references: EMBL:AL009204; PIDN:CAAL5793.1; GSPDB:GN00070; SCODEB:SC9B10.03c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCODEB:SC9B10.03c

C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase

Query Match 90.3%; Score 28; DB 2; Length 818;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 134 YTECR 138

RESULT 12

A27832

cell division control protein CDC16 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YKL022c

C:Species: Saccharomyces cerevisiae

C>Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 11-Jan-2000

C:Accession: A27832; S37839

R:Ichio, T.; Wickner, R.B.

Nucleic Acids Res. 15, 8439-8450, 1987

A:Title: Metal-binding, nucleic acid-binding finger sequences in the CDC16 gene of Sacch

A:Reference number: A27832; MUID:88040465; PMID:2823230

A:Accession: A27832

A:Molecule type: DNA

A:Residues: 1-840 <ICHS>

A:Cross-references: EMBL:X06165; NID:g3477; PIDN:CAAL29521.1; PID:g3478

R:Reger, M.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S37832

A:Accession: S37839

A:Molecule type: DNA

A:Residues: 1-840 <RIS>

A:Cross-references: EMBL:Z28022; NID:g486014; PIDN:CAAB1857.1; PID:g486015; MIPS:YKL022c

A:Experimental source: strain S286C

C:Genetics:

A:Gene: SGD:CDIC16

A:Cross-references: SGD:S0001505; MIPS:YKL022c

A:Map position: 11L

C:Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat ho
C:Keywords: cell cycle control; nucleus
F:56-489/Domain: tetratricopeptide repeat homology <TT1>
F:497-530/Domain: tetratricopeptide repeat homology <TT2>
F:531-564/Domain: tetratricopeptide repeat homology <TT3>
F:565-598/Domain: tetratricopeptide repeat homology <TT4>

F:599-632/Domain: tetratricopeptide repeat homology <TT5>
F:633-666/Domain: tetratricopeptide repeat homology <TT6>

Query Match 90.3%; Score 28; DB 2; Length 840;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 474 YTOCK 478

RESULT 13

AB1914

DNA gyrase A chain [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AB1914

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 6, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB1914

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-872 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW72817.1; PID:g17130205; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0860

C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomeras

Query Match 90.3%; Score 28; DB 2; Length 872;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 122 YTECR 126

RESULT 14

AF3362

DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain A [imported] - Brucella melit

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Jun-2002

C:Accession: AF3362

R:DeLVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mufer, C.; Los, T.; Ivanov

; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit

A:Reference number: AD3252; PMID:11756688

A:Accession: AF3362

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-902 <KUR>

A:Cross-references: GB:AF008917; PIDN:AAL52065.1; PID:g17982834; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME10884

A:Map position: 1

C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomeras

Query Match 90.3%; Score 28; DB 2; Length 902;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
Db 106 YTECR 110

RESULT 15

AB2762

DNA gyrase subunit A [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AB2762

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AB2762

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-906 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAI42512.1; PID:917739931; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: gyrA

A:Map position: circular chromosome

C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase

Query Match

Best Local Similarity 90.3%; Score 28; DB 2; Length 906;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YTOCR 5

DB 106 YTECR 110

Search completed: July 9, 2003, 12:02:46
Job time : 10.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:55:50 ; Search time 7.75 Seconds

(Without alignments)
26.759 Million cell updates/sec

Title: US-09-851-058-3

Perfect score: 31
Sequence: 1 YTCR 5

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	100.0	155 1	ECP1_MOUSE
2	31	100.0	155 1	ECP_RAT
3	31	100.0	156 1	ECP2_MOUSE
4	31	100.0	156 1	ECP3_MOUSE
5	28	90.3	304 1	PAGO_SALTY
6	28	90.3	432 1	VZ21_MERTIA
7	28	90.3	471 1	TNAA_ECO57
8	28	90.3	472 1	TNAA_ECOLI
9	28	90.3	472 1	TNAA_VIBCH
10	28	90.3	840 1	CC16_YEAST
11	27	87.1	61 1	RS14_CLOPE
12	27	87.1	61 1	RS14_HELPY
13	27	87.1	61 1	RS14_HELPY
14	27	87.1	89 1	IT12_ARATH
15	27	87.1	89 1	IT13_ARATH
16	27	87.1	99 1	IT12_SINAL
17	27	87.1	122 1	RBS2_HYDNR
18	27	87.1	210 1	PSBP_PORPU
19	27	87.1	251 1	HXB4_FUGRU
20	27	87.1	342 1	CRTB_STRGR
21	27	87.1	347 1	CHM1_CHICK
22	27	87.1	451 1	ILK_CAVPO
23	27	87.1	452 1	ILK1_HUMAN
24	27	87.1	452 1	ILK2_HUMAN
25	27	87.1	452 1	ILK_MOUSE
26	27	87.1	717 1	MYO1_HUMAN
27	27	87.1	729 1	CUL6_CAEEL
28	27	87.1	1122 1	TERT_MOUSE
29	26	83.9	64 1	RL35_DREPA
30	26	83.9	84 1	RNS_GIRCA
31	26	83.9	87 1	RL37_SCHMA
32	26	83.9	100 1	YJPO_YEAST
33	26	83.9	145 1	Y586_HAEIN

34	26	83.9	155 1	ECP5_MOUSE	035292 mus musculus
35	26	83.9	233 1	VHEL_PMY	P20952 papaya mosa
36	26	83.9	249 1	PS72_DROME	024178 drosophila
37	26	83.9	256 1	TPS1_BRAHE	095066 brachydanio
38	26	83.9	274 1	VGLL_MCMVK	P52513 murine cyto
39	26	83.9	274 1	VGLL_MCMVS	P22514 murine cyto
40	26	83.9	348 1	MMOC_MERCA	P22868 methylcoccc
41	26	83.9	374 1	SVN_AERPE	099924 aeropyrum p
42	26	83.9	374 1	YDB1_SCHPO	010354 sechizosach
43	26	83.9	395 1	Y414_MERTIA	057855 methanococ
44	26	83.9	440 1	PS4_CHICK	090732 gallus gall
45	26	83.9	440 1	PS4_HUMAN	003527 homo sapien

ALIGNMENTS

RESULT 1	ID	ECPL_MOUSE	STANDARD:	PRT:	155 AA.
AC	P97426:				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	Eosinophil cationic protein 1 precursor (EC 3.1.27.-) (ECP 1)				
DE	(Ribonuclease 3-1) (Rnase 3-1) (Eosinophil secondary granule				
DE	ribonuclease-1) (EAR-1).				
GN	EAR1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_Taxid=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J;				
RX	MEDLINE=97057249; PubMed=8901588;				
RA	Larson K.A., Olson E.V., Madden B.J., Gietz G.J., Lee N.A., Lee J.J.;				
RT	"Two highly homologous ribonuclease genes expressed in mouse				
RT	eosinophils identify a larger subgroup of the mammalian ribonuclease				
RT	superfamily."				
RL	Proc. Natl. Acad. Sci. U.S.A. 93:12370-12375(1996).				
CC	- FUNCTION: CYTOTOXIN AND HELMINTHOTOXIN WITH RIBONUCLEASE ACTIVITY.				
CC	- POSSESSES A WIDE VARIETY OF BIOLOGICAL ACTIVITIES (BY SIMILARITY).				
CC	- SUBCELLULAR LOCATION: MATRIX OF EOSINOPHIL'S LARGE SPECIFIC				
CC	GRANULE (BY SIMILARITY).				
CC	- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL: U72032; AAB37786.1; -				
DR	HSSP: P00656; 2RMS.				
DR	MGP: MGI-108021; Ear1.				
DR	InterPro: IPR001427; RNaseA.				
DR	Pfam: PF00074; RNaseA; 1.				
DR	PRINTS: PR00794; RIBONUCLEASE.				
DR	PRODOM: PD000535; RNaseA; 1.				
DR	SMART: SMO0092; RNase PC; 1.				
DR	PROSITE: PS00127; RNASE_PANCREATIC; 1.				
KW	Hydrolase; Nuclease; Endonuclease; Eosinophil; Toxin; Glycoprotein;				
FT	Signal.				
FT	SIGNAL	1	25	POTENTIAL.	
FT	CHAIN	26	155	EOSINOPHIL CATIONIC PROTEIN 1.	
FT	ACT_SITE	38	38	BY SIMILARITY.	
FT	ACT_SITE	62	62	BY SIMILARITY.	
FT	ACT_SITE	150	150	BY SIMILARITY.	
FT	DISULFID	47	106	BY SIMILARITY.	
FT	DISULFID	61	118	BY SIMILARITY.	

FT DISULFID 79 133 BY SIMILARITY.
 FT CARBOHYD 86 94 BY SIMILARITY.
 FT CARBOHYD 96 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 107 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 114 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 155 AA; 17296 MW; F8264E7A32B20D87 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 1; Length 155;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCOR 5
 DB 115 YTCOR 119

RESULT 2

ECP_RAT ID ECP_RAT STANDARD; PRT; 155 AA.
 AC P70709;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Eosinophil cationic protein precursor (EC 3.1.27.-) (ECP)
 DE (Ribonuclease 3) (Rnase 3).
 GN RNASE3 OR RN53.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Bone marrow;
 RX MEDLINE=97236431; PubMed=9116043;
 RA Nitoh T., Hirakata M., Mue S., Ohuchi K.;
 RT Identification of cDNA encoding rat eosinophil cationic
 RL protein/eosinophil-associated ribonuclease.";
 CC Blochem. Biophys. Acta 1351:42-46(1997).
 CC -1- FUNCTION: CYTOTOXIN AND HELMINTHOTOXIN WITH RIBONUCLEASE ACTIVITY.
 CC POSSESSES A WIDE VARIETY OF BIOLOGICAL ACTIVITIES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MATRIX OF EOSINOPHIL'S LARGE SPECIFIC
 CC GRANULE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D88586; BAA13648.1; -;
 DR HSSP; P00656; 2RNS.
 DR InterPro; IPR001427; RNaseA.
 DR Pfam; PF00074; rnasea.1.
 DR PRINTS; PR00794; RIBONUCLEASE.
 DR ProDom; PD000535; RNaseA.1.
 DR SMART; SM00092; RNase_Pc.1.
 DR PROSITE; PS00127; RNASE_PANCREATIC; 1.
 KW Hydrolyase; Nuclease; Endonuclease; Eosinophil; Toxin; Glycoprotein;
 KM Signal.
 FT SIGNAL. 1 25
 FT CHAIN 26 155 POTENTIAL.
 FT ACT_SITE 38 38 EOSINOPHIL CATIONIC PROTEIN.
 FT ACT_SITE 62 62 BY SIMILARITY.
 FT ACT_SITE 150 150 BY SIMILARITY.
 FT DISULFID 47 106 BY SIMILARITY.
 FT DISULFID 61 118 BY SIMILARITY.
 FT DISULFID 79 133 BY SIMILARITY.
 FT DISULFID 94 94 BY SIMILARITY.
 SQ SEQUENCE 155 AA; 18007 MW; F190864C34EA2C7 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 1; Length 155;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCOR 5
 DB 115 YTCOR 119

RESULT 3

ECP2_MOUSE ID ECP2_MOUSE STANDARD; PRT; 156 AA.
 AC P97425;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Eosinophil cationic protein 2 precursor (EC 3.1.27.-) (ECP 2)
 DE (Ribonuclease 3-2) (Rnase 3-2) (Eosinophil secondary granule
 DE ribonuclease-2) (EAR-2).
 GN EAR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=97057249; PubMed=8901588;
 RA Larson K.A., Olson E.V., Madden B.J., Gleich G.J., Lee N.A., Lee J.J.;
 RT "Two highly homologous ribonuclease genes expressed in mouse
 RT eosinophils identify a larger subgroup of the mammalian ribonuclease
 RT superfamily.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12370-12375(1996).
 CC -1- FUNCTION: CYTOTOXIN AND HELMINTHOTOXIN WITH RIBONUCLEASE ACTIVITY.
 CC POSSESSES A WIDE VARIETY OF BIOLOGICAL ACTIVITIES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MATRIX OF EOSINOPHIL'S LARGE SPECIFIC
 CC GRANULE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U72031; AAB37785.1; -;
 DR HSSP; P00656; 2RNS.
 DR MCD; MGI:108020; Ear2.
 DR InterPro; IPR001427; RNaseA.
 DR Pfam; PF00074; rnasea.1.
 DR PRINTS; PR00794; RIBONUCLEASE.
 DR ProDom; PD000535; RNaseA.1.
 DR SMART; SM00092; RNase_Pc.1.
 DR PROSITE; PS00127; RNASE_PANCREATIC; 1.
 KW Hydrolyase; Nuclease; Endonuclease; Eosinophil; Toxin; Glycoprotein;
 KM Signal.
 FT SIGNAL. 1 25
 FT CHAIN 26 156 POTENTIAL.
 FT ACT_SITE 38 38 EOSINOPHIL CATIONIC PROTEIN 2.
 FT ACT_SITE 62 62 BY SIMILARITY.
 FT ACT_SITE 151 151 BY SIMILARITY.
 FT ACT_SITE 151 151 BY SIMILARITY.
 FT DISULFID 47 106 BY SIMILARITY.
 FT DISULFID 61 119 BY SIMILARITY.
 FT DISULFID 79 134 BY SIMILARITY.
 FT DISULFID 86 94 BY SIMILARITY.
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 156 AA; 17620 MW; A158C1133091053 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 156;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
 |||||
 Db 116 YTOCR 120

RESULT 4

ECP3_MOUSE STANDARD; PRT; 156 AA.
 ID ECP3_MOUSE
 AC 035290;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Eosinophil cationic-type ribonuclease 3 precursor (MR-3).
 GN EAP3.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=97478550; PubMed=9336452;
 RA Batten D., Dyer K.D., Domachowski J.B., Rosenberg H.F.;
 RT "Molecular cloning of four novel murine ribonuclease genes: unusual
 expansion within the ribonuclease A gene family."
 RL Nucleic Acids Res. 25:4235-4239(1997).
 CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF017258; AAC53489.1; -
 CC HSSP: P00656; 2RNS.
 DR MGD: MGI:1858237; Ear3.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA; 1.
 DR PRINTS: PR00794; RIBONUCLEASE.
 DR ProDom: PD000535; RNaseA; 1.
 DR SMART: SM00092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNASE_PANCREATIC; 1.
 KW Hydrolase; Nuclease; Endonuclease; Glycoprotein; Signal.
 FT SIGNAL 1..25
 FT CHAIN 26..156
 FT ACT_SITE 38
 FT ACT_SITE 62
 FT ACT_SITE 151
 FT DISULFID 47
 FT DISULFID 61
 FT DISULFID 79
 FT DISULFID 86
 FT CARBOHYD 41
 FT CARBOHYD 89
 FT CARBOHYD 96
 FT CARBOHYD 107
 FT SEQUENCE 156 AA; 17744 MW; C9C2B6CEB5DA06C CRC64;

Query Match 100.0%; Score 31; DB 1; Length 156;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
 |||||
 Db 116 YTOCR 120

RESULT 5

PAGO_SALTY STANDARD; PRT; 304 AA.
 ID PAGO_SALTY
 AC 030646;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE PAGO protein.
 GN PAGO OR STM1862.
 OS Salmonella typhimurium.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 CC NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14028S;
 RX MEDLINE=98380520; PubMed=9712687;
 RA Gunn J.S., Belden W.J., Miller S.I.;
 RT "Identification of phop-phoQ activated genes within a duplicated
 RT region of the Salmonella typhimurium chromosome."
 RL Microb. Pathog. 25:77-90(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Lt2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Potvin J.K., All J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT Lt2."
 RL Nature 413:852-856(2001).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE EMA TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF013775; AAB82452.1; -
 CC EMBL: AE008783; AAL20777.1; -
 DR StyGene: SG10678; pago.
 DR InterPro: IPR000620; DUF6.
 DR Pfam: PF00892; DUF6; 2.
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 4..24
 FT TRANSMEM 34
 FT TRANSMEM 54
 FT TRANSMEM 67
 FT TRANSMEM 95
 FT TRANSMEM 119
 FT TRANSMEM 139
 FT TRANSMEM 150
 FT TRANSMEM 180
 FT TRANSMEM 200
 FT TRANSMEM 214
 FT TRANSMEM 246
 FT TRANSMEM 267
 FT SEQUENCE 304 AA; 33622 MW; B368153F6EBFB0B CRC64;

Query Match 90.3%; Score 28; DB 1; Length 304;
 Best Local Similarity 80.0%; Pred. No. 51;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTOCR 5
 |||||
 Db 167 YTOCR 171

RESULT 6

Y221_METUA
ID Y221_METUA STANDARD; PRT; 432 AA.
AC 060281;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJECL21.
GN MJECL21.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota: Methanococci: Methanococcales:
OC Methanocaldococcaceae: Methanocaldococcus:
OX NCBI_TaxID=2190;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RA MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierulff A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeck R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Feldman J.F., Fuhrmann J.L., Nguyen D.,
RA Utechtack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL: L77118; AAC37092.1;
DR TIGR: MJECL21;
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 432 AA; 51081 MW; DBADP2CSC43M4F90 CRC64;
Query Match 90.3%; Score 28; DB 1; Length 432;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YTCR 5
DB 299 YTECR 303

RESULT 7
TNAI_ECO57 STANDARD; PRT; 471 AA.
ID TNAI_ECO57
AC 08X834;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
GN TNAI OR 25203 OR ECS4645.
OS Escherichia coli O157:H7.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
OC Escherichia.
OX NCBI_TaxID=83334;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RA MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouls K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RA MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC
CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O = indole + pyruvate +
CC NH(3).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Tryptophan catabolism.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL: AE005602; AAG58908.1; ALT_INT.
DR EMBL: AP002566; BAB38068.1; ALT_INT.
DR InterPro: IPR001597; Beta_elim_lyase.
DR Pfam: PF01212; Beta_elim_lyase.1.
DR ProDom: PD005927; Beta_elim_lyase.1.
DR PROSITE: PS00853; BETA_ELIM_LYASE.1.
KW Tryptophan catabolism; Lyase; Pyridoxal phosphate; Complete proteome.
FT BINDING 270 270
SQ SEQUENCE 471 AA; 52789 MW; 49DD2D41BD9D0034 CRC64;
Query Match 90.3%; Score 28; DB 1; Length 471;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YTCR 5
DB 291 YTECR 295

RESULT 8
TNAI_ECOLI STANDARD; PRT; 471 AA.
ID TNAI_ECOLI
AC P00913; P78123;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
GN TNAI OR IND OR B5708.
OS Escherichia coli.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
OC Escherichia.
OX NCBI_TaxID=562;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA MEDLINE=82007678; PubMed=6268608;
RA Deeley M.C., Yanofsky C.;
RT "Nucleotide sequence of the structural gene for tryptophanase of
RT Escherichia coli K-12.";
RL J. Bacteriol. 147:787-796(1981).
RN (2)
RP SEQUENCE FROM N.A., AND IMPORTANCE OF CYS-298.
RC STRAIN=B/177-A;
RA MEDLINE=89323226; PubMed=2502187;
RA Tokushige M., Tsujimoto N., Oda T., Yumoto N., Ito S.;

RA Yamamoto M., Kim E.H., Hiragi Y.;
 RT "Role of cysteine residues in tryptophanase for monovalent cation-
 induced activation.";
 RL Biochimie 71:711-720(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-93315143; PubMed-7686882;
 RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
 RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
 genome: organizational symmetry around the origin of replication.";
 RL Genomics 16:551-561(1993).
 RN [4]
 RP SEQUENCE OF TRYPTIC PEPTIDES.
 RC STRAIN-K12;
 RX MEDLINE-72134434; PubMed-4551944;
 RA Kagamiyama H., Matsubara H., Sneli E.E.;
 RT "The chemical structure of tryptophanase from Escherichia coli. 3.
 RT Isolation and amino acid sequence of the tryptic peptides.";
 RL J. Biol. Chem. 247:1576-1586(1972).
 RN [5]
 RP SEQUENCE OF 1-21 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-86033634; PubMed-3902796;
 RA Stewart V., Yanofsky C.;
 RT "Evidence for transcription antitermination control of tryptophanase
 RT operon expression in Escherichia coli K-12.";
 RL J. Bacteriol. 164:731-740(1985).
 RN [6]
 RP SEQUENCE OF 463-471 FROM N.A.
 RX MEDLINE-91216998; PubMed-2022620;
 RA Sarsero J.P., Wookley P.J., Gollnick P.D., Yanofsky C., Pittard A.J.;
 RT "A new family of integral membrane proteins involved in transport of
 RT aromatic amino acids in Escherichia coli.";
 RL J. Bacteriol. 173:3231-3234(1991).
 RN [7]
 RP SEQUENCE OF 1-12.
 RC STRAIN-K12 / EMG2;
 RX MEDLINE-97443975; PubMed-9298646;
 RA Link A.J., Robinson K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of Escherichia coli K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 RN [8]
 RP MUTAGENESIS OF CYS-294 AND CYS-298.
 RX MEDLINE-89278130; PubMed-2659590;
 RA Phillips R.S., Gollnick P.D.;
 RT "Evidence that cysteine 298 is in the active site of tryptophan
 RT indole-lyase.";
 RL J. Biol. Chem. 264:10627-10632(1989).
 CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> indole + pyruvate +
 CC NH(3).
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- PATHWAY: Tryptophan catabolism.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: R00032; AAA24676.1; -;
 CC EMBL: X15974; CAA34096.1; -;
 CC EMBL: M11990; AAA24679.1; -;
 CC EMBL: M59914; -; NOT ANNOTATED CDS.
 CC EMBL: L10328; AAA62059.1; ALT_INIT.
 CC EMBL: AE000446; AAC76731.1; ALT_INIT.
 CC PIR: A01136; WZEC.
 CC HSP: P28796; IAX4.

DR ECO2DBASE; G046.5; 6TH EDITION.
 DR Ecogene; EG11005; tnaA.
 DR InterPro; IPR001597; Beta_elim_lyase.
 DR Pfam; PF01212; beta_elim_lyase.1.
 DR ProDom; PD005927; Beta_elim_lyase.1.
 DR ProSite; PS00853; BETA_ELIM_LYASE.
 DR TRYPPTOPHAN catabolism; Lyase; PYRIDOXAL PHOSPHATE.
 FT BINDING 270 270
 FT MUTAGEN 294 294 C->S: IDENTICAL TO WILD-TYPE.
 FT CONFLICT 137 140 DFTQ -> TTGG (IN REF. 1).
 FT CONFLICT 379 380 QA -> TG (IN REF. 1 AND 2).
 SQ SEQUENCE 471 AA; 52773 MW; 5AFC1F41BD9D0034 CRC64;
 Query Match 90.3%; Score 28; DB 1; Length 471;
 Best Local Similarity 80.0%; Pred. No. 75;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 DB 291 YTECR 295
 QY 1 YTCOR 5
 DB 11:11
 ID TNA_A_VIBCH STANDARD; PRT; 472 AA.
 AC O9KN05; 086039;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
 GN TNA OR YCA0161.
 OS Vibrio cholerae.
 CC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=66;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-El Tor N16961 / Serotype O1;
 RX MEDLINE-20406833; PubMed-10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwim M.L.,
 RA Dodson R.D., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettlin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., McKelano J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 RN [2]
 RP SEQUENCE OF 51-241 FROM N.A.
 RC STRAIN-Bengali;
 RA Tang T.H., Ravichandran M., Johari M.R., Zalanuddin Z.F.;
 RT "Vibrio cholerae putative tryptophanase gene partial cds.";
 RL Submitted (Jul-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O -> indole + pyruvate +
 CC NH(3).
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -1- PATHWAY: Tryptophan catabolism.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE004357; AAF96074.1; -;
 CC EMBL: AF081274; AAC33284.1; -;
 CC HSP: P28796; IAX4.

DR TIGR: VCA0161: -
 DR InterPro: IPR001597; Beta_alm_lyase.
 DR Pfam: PF01212; Beta_alm_lyase; 1.
 DR ProDom: PD005927; Beta_alm_lyase; 1.
 DR PROSITE: PS00853; BETA_ALM_LYASE; 1.
 DR TrpTophan catabolism: Lyase: Pyridoxal phosphate: Complete proteome.
 FT BINDING 270 270
 SQ SEQUENCE 472 AA; 52920 MW; 632384F54C06F7B9 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 472;
 Best Local Similarity 80.0%; Pred. No. 75;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTCOR 5
 Db 291 YTECR 295

RESULT 10
 CC16_YEAST STANDARD; PRT; 840 AA.
 ID CC16_YEAST
 AC P09798;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Cell division control protein 16.
 GN CDC16 OR YKL022C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=R.B. Wickner 1385;
 RX MEDLINE=88040465; PubMed=2823230;
 RA Icho T., Wickner R.B.;
 RT "Metal-binding, nucleic acid-binding finger sequences in the CDC16
 RT gene of *Saccharomyces cerevisiae*.";
 RL Nucleic Acids Res. 15:8439-8450(1987).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC Rieger M.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP TPR REPEATS.
 RX MEDLINE=90124639; PubMed=2404612;
 RA Sikorski R.S., Boguski M.S., Goebel M., Hieter P.A.;
 RT "A repeating amino acid motif in CDC23 defines a family of proteins
 RT and a new relationship among genes required for mitosis and RNA
 RT synthesis.";
 RL Cell 60:307-317(1990).
 RN (4)
 RP SUBUNITS.
 RX MEDLINE=95009933; PubMed=7925276;
 RA Lamb J.R., Michaud W.A., Sikorski R.S., Hieter P.A.;
 RT "cdc16p, cdc23p and cdc27p form a complex essential for mitosis.";
 RL EMBO J. 13:4321-4328(1994)
 CC -1- FUNCTION: EXACT FUNCTION NOT KNOWN. REQUIRED FOR CHROMOSOME
 CC SEGREGATION. MUTATIONS IN CDC16 CAUSE CELLS TO ARREST UNIFORMELY
 CC AT G2/M AFTER DNA REPLICATION BUT PRIOR TO MITOTIC SPINDLE
 CC ELONGATION.
 CC -1- SUBUNIT: CDC16, CDC23 AND CDC27 FORMS A MACROMOLECULAR COMPLEX.
 CC -1- SIMILARITY: CONTAINS 10 TPR REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X06165; CA429521.1;
 DR EMBL: Z28022; CA481857.1;
 DR PIR: A27832; A27832
 DR SGD: S0001505; CDC16.
 DR InterPro: IPR001440; TPR.
 DR Pfam: PF00515; TPR; 7.
 DR SMART: SM00028; TPR; 6.
 KW Cell division; Cell cycle; Mitosis; Repeat; TPR repeat;
 KW Nuclear protein.
 FT REPEAT 263 295 TPR 1.
 FT REPEAT 296 329 TPR 2.
 FT REPEAT 359 392 TPR 3.
 FT REPEAT 497 530 TPR 4.
 FT REPEAT 531 564 TPR 5.
 FT REPEAT 565 598 TPR 6.
 FT REPEAT 599 632 TPR 7.
 FT REPEAT 633 666 TPR 8.
 FT REPEAT 674 707 TPR 9.
 FT REPEAT 708 741 TPR 10.
 SQ SEQUENCE 840 AA; 94991 MW; A096B34441083488 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 840;
 Best Local Similarity 80.0%; Pred. No. 1,2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTCOR 5
 Db 474 YTOCK 478

RESULT 11
 RS14_CLOPE STANDARD; PRT; 61 AA.
 ID RS14_CLOPE
 AC ORXHT6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S14.
 GN RPSN OR CPE2392.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 RX NCBI_TaxID=1502;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / Type A;
 RX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of *Clostridium perfringens*, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 CC -1- FUNCTION: Known to be required for the assembly of 30S particles
 CC and may also be responsible for determining the conformation of
 CC the 16S RNA at the A site (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

Query Match 87.1%; Score 27; DB 1; Length 61;

Best Local Similarity 80.0%; Pred. No. 20;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
11:11
DB 21 YTCR 25

RESULT 12

RS14_HELPJ STANDARD: PRT: 61 AA.

ID RS14_HELPJ

AC 092353;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 30S ribosomal protein S14.

GN RPSN OR HP1226.

OS Helicobacter pylori J99 (Campylobacter pylori J99).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

CC Helicobacter.

OX NCBI_TaxID=85963;

RP [1]

SEQUENCE FROM N.A.

RA MEDLINE-99120557; PubMed-9923682;

RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doly P.C.,

RA Smith D.R., Noonan B., Guild B.C., deJonghe B.L., Carmel G.,

RA Tummino P.J., Caruso A., Urie-Nickelsen M., Mills D.M., Ives C.,

RA Gibson R., Meberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,

RA Trust T.J.;

RT "Genomic sequence comparison of two unrelated isolates of the human

RT gastric pathogen Helicobacter pylori.";

RL Nature 397:176-180(1999).

CC -1- FUNCTION: KNOWN TO BE REQUIRED FOR THE ASSEMBLY OF 30S PARTICLES

CC AND MAY ALSO BE RESPONSIBLE FOR DETERMINING THE CONFORMATION OF

CC THE 16S RNA AT THE A SITE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE001547; AAD06792.1;

DR InterPro: IPR001209; Ribosomal_S14.

DR Pfam: PF00253; Ribosomal_S14; 1.

DR PROSITE: PS00527; RIBOSOMAL_S14; 1.

KW Ribosomal protein; Complete proteome.

SO SEQUENCE 61 AA; 7056 MW; EB4D764FCF57EAB6 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 61;

Best Local Similarity 80.0%; Pred. No. 20;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5

11:11

DB 21 YTCR 25

RESULT 13

RS14_HELPJ STANDARD: PRT: 61 AA.

ID RS14_HELPJ

AC P56021;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 30S ribosomal protein S14.

GN RPSN OR HP1306.

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

CC Helicobacter.

OX NCBI_TaxID=210;

RP [1]

SEQUENCE FROM N.A.

RA STRAIN-26595 / ATCC 700392;

RA MEDLINE-97394467; PubMed-9252185;

RA Tomb J.-F., White O., Keriavage A.R., Clayton R.A., Sutton G.G.,

RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,

RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,

RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,

RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.R.,

RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Wathey L., Wallin E.,

RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,

RA Venter J.C.;

RT "The complete genome sequence of the gastric pathogen Helicobacter

RT pylori.";

RL Nature 388:539-547(1997).

CC -1- FUNCTION: KNOWN TO BE REQUIRED FOR THE ASSEMBLY OF 30S PARTICLES

CC AND MAY ALSO BE RESPONSIBLE FOR DETERMINING THE CONFORMATION OF

CC THE 16S RNA AT THE A SITE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE000633; AAD08361.1;

DR TIGR: HP1306;

DR InterPro: IPR001209; Ribosomal_S14.

DR Pfam: PF00253; Ribosomal_S14; 1.

DR PROSITE: PS00527; RIBOSOMAL_S14; 1.

KW Ribosomal protein; Complete proteome.

SO SEQUENCE 61 AA; 7042 MW; EB4D764FCF507AB6 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 61;

Best Local Similarity 80.0%; Pred. No. 20;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5

11:11

DB 21 YTCR 25

RESULT 14

IT12_ARATH STANDARD: PRT: 89 AA.

ID IT12_ARATH

AC 042328;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Putative trypsin inhibitor AT1-2 precursor.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RP [1]

SEQUENCE FROM N.A.

RA STRAIN-CV, Columbia;

RT "Nucleotide sequence of an Arabidopsis cDNA encoding a serine

RT proteinase inhibitor AT1-2.";

RL (in) Plant Gene Register PCR95-011.

RN [2]

RP CONCEPTUAL TRANSLATION.

RA Batroch A.;

Unpublished observations (JUN-1998).

CC -1- SIMILARITY: BELONGS TO THE RT1/MT1-2 PROTEASE INHIBITORS FAMILY.

CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; FRAMESHIFTS WERE

CC ----- CORRECTED IN THE SIGNAL PEPTIDE REGION. -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC of the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Z46816; CAA6849.1; ALT_FRAME.

DR InterPro: IPR003614; Knc01.

DR SMART: SM00505; Knc01: 1.

KW Serine protease inhibitor; Signal; Multigene family.

FT CHAIN 1 27 POTENTIAL.

FT DISULFID 28 89 PUTATIVE TRYPSIN INHIBITOR ATG43510.

FT DISULFID 32 84 BY SIMILARITY.

FT DISULFID 45 69 BY SIMILARITY.

FT DISULFID 54 79 BY SIMILARITY.

FT DISULFID 58 81 BY SIMILARITY.

FT ACT_SITE 48 49 REACTIVE-BOND (BY SIMILARITY).

SO SEQUENCE 89 AA; 9911 MW; 35ECFD89E944E099 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 89;

Best Local Similarity 80.0%; Pred. No. 28;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5

Db 55 YTCR 59

RESULT 15

ID ITI3_ARATH STANDARD; PRT; 89 AA.

AC 022865;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative trypsin inhibitor At2g43510 precursor (dId1 4T-1).

CN ATG43510 OR T01024.25.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Vercauteren I.J.R., Van Der Schueren E., Van Montagu M., Cheysen G.;

RT "Molecular investigation of nematode-induced galls in Arabidopsis

RL thaliana using the differential display technique."

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA MEDLINE=20083487; Pubmed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Bernstead M.E., Feldlyum T.V.,

RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H.L.,

RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,

RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,

RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,

RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,

RA Venter J.C.;

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis

RT thaliana";

RL Nature 402:761-768(1999).

CC -1- SIMILARITY: BELONGS TO THE RTI/MTI-2 PROTEASE INHIBITORS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AJ249958; CAB62548.1;

DR EMBL: AC002335; AAB64325.1;

DR InterPro: IPR003614; Knc01.

DR SMART: SM00505; Knc01: 1.

KW Serine protease inhibitor; Signal; Multigene family.

FT SIGNAL 1 27 POTENTIAL.

FT CHAIN 28 89 PUTATIVE TRYPSIN INHIBITOR ATG43510.

FT DISULFID 32 84 BY SIMILARITY.

FT DISULFID 45 69 BY SIMILARITY.

FT DISULFID 54 79 BY SIMILARITY.

FT DISULFID 58 81 BY SIMILARITY.

FT ACT_SITE 48 49 REACTIVE-BOND (BY SIMILARITY).

SO SEQUENCE 89 AA; 9885 MW; C4FA5D0B3ABB03D5 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 89;

Best Local Similarity 80.0%; Pred. No. 28;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTCR 5

Db 55 YTCR 59

Search completed: July 9, 2003, 12:00:45

Job time : 8.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:57:21 ; Search time 17.75 Seconds

(Without alignments)
58.041 Million cell updates/sec

Title: US-09-851-058-3

Perfect score: 31

Sequence: 1 YNGCR 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rhodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	31	100.0	153	11	Q9R130	Q9R130 ratu
2	31	100.0	153	11	Q9R130	Q9R130 ratu
3	31	100.0	153	11	Q9R130	Q9R130 ratu
4	31	100.0	153	11	Q9R130	Q9R130 ratu
5	31	100.0	153	11	Q9R130	Q9R130 ratu
6	31	100.0	153	11	Q9R130	Q9R130 ratu
7	31	100.0	153	11	Q9R130	Q9R130 ratu
8	31	100.0	153	11	Q9R130	Q9R130 ratu
9	31	100.0	153	11	Q9R130	Q9R130 ratu
10	31	100.0	153	11	Q9R130	Q9R130 ratu
11	31	100.0	153	11	Q9R130	Q9R130 ratu
12	31	100.0	153	11	Q9R130	Q9R130 ratu
13	31	100.0	153	11	Q9R130	Q9R130 ratu
14	31	100.0	153	11	Q9R130	Q9R130 ratu
15	31	100.0	153	11	Q9R130	Q9R130 ratu
16	31	100.0	153	11	Q9R130	Q9R130 ratu

17	31	100.0	153	11	Q9R130	Q9R130 ratu
18	31	100.0	153	11	Q9R130	Q9R130 ratu
19	31	100.0	153	11	Q9R130	Q9R130 ratu
20	31	100.0	153	11	Q9R130	Q9R130 ratu
21	31	100.0	153	11	Q9R130	Q9R130 ratu
22	31	100.0	153	11	Q9R130	Q9R130 ratu
23	31	100.0	153	11	Q9R130	Q9R130 ratu
24	31	100.0	153	11	Q9R130	Q9R130 ratu
25	31	100.0	153	11	Q9R130	Q9R130 ratu
26	31	100.0	153	11	Q9R130	Q9R130 ratu
27	31	100.0	153	11	Q9R130	Q9R130 ratu
28	31	100.0	153	11	Q9R130	Q9R130 ratu
29	31	100.0	153	11	Q9R130	Q9R130 ratu
30	31	100.0	153	11	Q9R130	Q9R130 ratu
31	31	100.0	153	11	Q9R130	Q9R130 ratu
32	31	100.0	153	11	Q9R130	Q9R130 ratu
33	31	100.0	153	11	Q9R130	Q9R130 ratu
34	31	100.0	153	11	Q9R130	Q9R130 ratu
35	31	100.0	153	11	Q9R130	Q9R130 ratu
36	31	100.0	153	11	Q9R130	Q9R130 ratu
37	31	100.0	153	11	Q9R130	Q9R130 ratu
38	31	100.0	153	11	Q9R130	Q9R130 ratu
39	31	100.0	153	11	Q9R130	Q9R130 ratu
40	31	100.0	153	11	Q9R130	Q9R130 ratu
41	31	100.0	153	11	Q9R130	Q9R130 ratu
42	31	100.0	153	11	Q9R130	Q9R130 ratu
43	31	100.0	153	11	Q9R130	Q9R130 ratu
44	31	100.0	153	11	Q9R130	Q9R130 ratu
45	31	100.0	153	11	Q9R130	Q9R130 ratu

ALIGNMENTS

RESULT 1
ID Q9R130 PRELIMINARY; PRT; 153 AA.
AC Q9R130;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribonuclease 7 precursor (Fragment).
GN R7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20063779; PubMed-10594173;
RA Slingshanta N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RA Domachowski J.B., Rosenberg H.F.;
RT of independent gene clusters in rats and mice.*;
RL J. Mol. Evol. 49:721-728(1999).
DR EMBL: AF171645; AAD51665.1; -;
DR HSP; P00656; 2RNS.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA; 1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR PRODOM: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase_Pc; 1.
DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 25
FT CHAIN 26 >153 POTENTIAL.
FT NON_TER 153 RIBONUCLEASE 7.
SQ SEQUENCE 153 AA; 17679 MW; B5878251BA734D6B CRC64;
Query Match 100.0%; Score 31; DB 11; Length 153;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 113 YTOCR 117

RESULT 2

09UK17 PRELIMINARY; PRT: 153 AA.
AC 09UK17;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Eosinophil-associated ribonuclease 4.
GN EARA4.
OS Mus saxicola (Spiny mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20243759; PubMed=10758160;
RA Zhang J., Dyer K.D., Rosenberg H.F.;
RT Evolution of the rodent eosinophil-associated ribonuclease gene family by rapid gene sorting and positive selection.*;
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
DR EMBL; AF238396; AAF67696.1; -.
DR HSP: P00656; 2RNS.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA; 1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR ProDom: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase_Pc; 1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
SQ SEQUENCE 153 AA; 17042 MW; CFF982F0A6FE296 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 11; Length 153;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 115 YTOCR 119

RESULT 3

09R134 PRELIMINARY; PRT: 155 AA.
AC 09R134;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribonuclease 1 precursor (Fragment).
GN R1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20063779; PubMed=10594173;
RA Singhania N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RT Domachowski J.B., Rosenberg H.F.;
RT Rapid evolution of the ribonuclease A superfamily: adaptive expansion of independent gene clusters in rats and mice.*;
RL J. Mol. Evol. 49:721-728(1999).
DR EMBL; AF171641; AAD51661.1; -.
DR HSP: P00656; 2RNS.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA; 1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR ProDom: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase_Pc; 1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
RN [1]

KW Signal.
FT SIGNAL. 1 25
FT CHAIN 26 >155
FT NON_TER 155
SQ SEQUENCE 155 AA; 18119 MW; 16BAA465D3260CEB CRC64;
Query Match
Best Local Similarity 100.0%; Score 31; DB 11; Length 155;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 115 YTOCR 119

RESULT 4

09R132 PRELIMINARY; PRT: 155 AA.
AC 09R132;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribonuclease 4 precursor (Fragment).
GN R4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20063779; PubMed=10594173;
RA Singhania N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RT Domachowski J.B., Rosenberg H.F.;
RT Rapid evolution of the ribonuclease A superfamily: adaptive expansion of independent gene clusters in rats and mice.*;
RL J. Mol. Evol. 49:721-728(1999).
DR EMBL; AF171643; AAD51663.1; -.
DR HSP: P00656; 1RBD.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA; 1.
DR PRINTS: PR00794; RIBONUCLEASE.
DR ProDom: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase_Pc; 1.
DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
RN [1]

Query Match
Best Local Similarity 100.0%; Score 31; DB 11; Length 155;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YTOCR 5
Db 115 YTOCR 119

RESULT 5

09R131 PRELIMINARY; PRT: 155 AA.
AC 09R131;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribonuclease 5 precursor (Fragment).
GN R5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-20063779; PubMed-10594173;
 RA Slingshania N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
 RA Domachowski J.B., Rosenberg H.F.;
 RT "Rapid evolution of the ribonuclease A superfamily: adaptive expansion
 RT of independent gene clusters in rats and mice."
 RL J. Mol. Evol. 49:721-728(1999).
 DR EMBL: AF171644; AAD51664.1; -.
 DR HSSP: P00656; IRD.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA; 1.
 DR PRINTS: PR00794; RIBONUCLEASE.
 DR PRODOM: PD000535; RNaseA; 1.
 DR SMART: SM00092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.
 KW Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 >155 RIBONUCLEASE 5.
 FT NON_TER 155 155
 SQ SEQUENCE 155 AA; 17189 MW; 36C6F3381DB92787 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 155;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
 Db 115 YTCR 119

RESULT 6
 09R127 PRELIMINARY; PRT; 155 AA.
 ID 09R127
 AC 09R127
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Ribonuclease 14 precursor (Fragment).
 GN R14.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID-10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20063779; PubMed-10594173;
 RA Slingshania N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
 RA Domachowski J.B., Rosenberg H.F.;
 RT "Rapid evolution of the ribonuclease A superfamily: adaptive expansion
 RT of independent gene clusters in rats and mice."
 RL J. Mol. Evol. 49:721-728(1999).
 DR EMBL: AF171648; AAD51668.1; -.
 DR HSSP: P00656; IRD.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA; 1.
 DR PRINTS: PR00794; RIBONUCLEASE.
 DR PRODOM: PD000535; RNaseA; 1.
 DR SMART: SM00092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.
 KW Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 >155 RIBONUCLEASE 14.
 FT NON_TER 155 155
 SQ SEQUENCE 155 AA; 17332 MW; 4E7B83380A0C68 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 155;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 YTCR 5
 Db 115 YTCR 119

RESULT 7
 09R125 PRELIMINARY; PRT; 155 AA.
 ID 09R125
 AC 09R125
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Ribonuclease 8 precursor (Fragment).
 GN EAR8 OR R8.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID-10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20063779; PubMed-10594173;
 RA Slingshania N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
 RA Domachowski J.B., Rosenberg H.F.;
 RT "Rapid evolution of the ribonuclease A superfamily: adaptive expansion
 RT of independent gene clusters in rats and mice."
 RL J. Mol. Evol. 49:721-728(1999).
 DR EMBL: AF171650; AAD51670.1; -.
 DR HSSP: P00656; 2RNS.
 DR MCD: MGI:1858206; Bar8.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA; 1.
 DR PRINTS: PR00794; RIBONUCLEASE.
 DR PRODOM: PD000535; RNaseA; 1.
 DR SMART: SM00092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.
 KW Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 >155 RIBONUCLEASE 8.
 FT NON_TER 155 155
 SQ SEQUENCE 155 AA; 17364 MW; D5B4DB5129D6AE2B CRC64;

Query Match 100.0%; Score 31; DB 11; Length 155;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
 Db 115 YTCR 119

RESULT 8
 09J16 PRELIMINARY; PRT; 155 AA.
 ID 09J16
 AC 09J16
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Eosinophil-associated ribonuclease 6.
 GN EAR6.
 OS Mus saxicola (Spiny mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID-10094;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20243759; PubMed-10758160;
 RA Zhang J., Dyer K.D., Rosenberg H.F.;
 RT "Evolution of the rodent eosinophil-associated ribonuclease gene
 RT family by rapid gene sorting and positive selection."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
 DR EMBL: AF238397; AAF67697.1; -.
 DR HSSP: P00656; 1LSO.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA; 1.
 DR PRINTS: PR00794; RIBONUCLEASE.
 DR PRODOM: PD000535; RNaseA; 1.
 DR SMART: SM00092; RNase_Pc; 1.

DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
 SO SEQUENCE 155 AA; 17363 MW; 8B5401F0C6CA1EF CRC64;

Query Match 100.0%; Score 31; DB 11; Length 155;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
 Db 115 YTCR 119

RESULT 9

ID O9JKI4 PRELIMINARY; PRT; 155 AA.

AC O9JKI4; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Eosinophil-associated ribonuclease 11.
 GN EAR11.
 OS Mus saxicola (Spiny mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10094;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-20243759; PubMed-10758160;
 RT Zhang J., Dyer K.D., Rosenberg H.F.;
 "Evolution of the rodent eosinophil-associated ribonuclease gene
 family by rapid gene sorting and positive selection."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
 DR EMBL; AF238399; AAF67699.1;
 DR HSSP; P00656; IL50.
 DR InterPro; IPR001427; RNaseA.
 DR Pfam; PF00074; RNaseA; 1.
 DR PRINTS; PR00794; RIBONUCLEASE.
 DR ProDom; PD000535; RNaseA; 1.
 DR SMART; SM00092; RNase_Pc; 1.
 DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
 SO SEQUENCE 155 AA; 17268 MW; D67163B8B0564A9 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 155;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
 Db 115 YTCR 119

RESULT 10

ID O9JKI2 PRELIMINARY; PRT; 155 AA.

AC O9JKI2; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Eosinophil-associated ribonuclease 23.
 GN FAR23.
 OS Mus saxicola (Spiny mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10094;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-20243759; PubMed-10758160;
 RT Zhang J., Dyer K.D., Rosenberg H.F.;
 "Evolution of the rodent eosinophil-associated ribonuclease gene
 family by rapid gene sorting and positive selection."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
 DR EMBL; AF238401; AAF67701.1;
 DR HSSP; P00656; IL50.

DR InterPro; IPR001427; RNaseA.
 DR Pfam; PF00074; RNaseA; 1.
 DR PRINTS; PR00794; RIBONUCLEASE.
 DR ProDom; PD000535; RNaseA; 1.
 DR SMART; SM00092; RNase_Pc; 1.
 DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
 SO SEQUENCE 155 AA; 17367 MW; 5B4F44D97B0C5E25 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 155;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
 Db 115 YTCR 119

RESULT 11

ID O9JKI0 PRELIMINARY; PRT; 155 AA.

AC O9JKI0; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Eosinophil-associated ribonuclease 7.
 GN EAR7.
 OS Mus pahari (Shrew mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10093;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-20243759; PubMed-10758160;
 RT Zhang J., Dyer K.D., Rosenberg H.F.;
 "Evolution of the rodent eosinophil-associated ribonuclease gene
 family by rapid gene sorting and positive selection."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
 DR EMBL; AF238403; AAF67703.1;
 DR HSSP; P00656; IL50.
 DR InterPro; IPR001427; RNaseA.
 DR Pfam; PF00074; RNaseA; 1.
 DR PRINTS; PR00794; RIBONUCLEASE.
 DR ProDom; PD000535; RNaseA; 1.
 DR SMART; SM00092; RNase_Pc; 1.
 DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
 SO SEQUENCE 155 AA; 17095 MW; 9086C9F9A0C8498F CRC64;

Query Match 100.0%; Score 31; DB 11; Length 155;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTCR 5
 Db 115 YTCR 119

RESULT 12

ID O9JKH9 PRELIMINARY; PRT; 155 AA.

AC O9JKH9; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Eosinophil-associated ribonuclease 8.
 GN EAR8.
 OS Mus pahari (Shrew mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10093;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-20243759; PubMed-10758160;
 RT Zhang J., Dyer K.D., Rosenberg H.F.;

```
RT "Evolution of the rodent eosinophil-associated ribonuclease gene  
family by rapid gene sorting and positive selection.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).  
DR EMBL: AF238404; AAF67704.1; -.  
DR HSSP: P00656; 2RNS.  
DR InterPro: IPR001427; RNaseA.  
DR Pfam: PF000794; RNaseA; 1.  
DR PRINTS: PR00794; RIBONUCLEASE.  
DR ProDom: PD000535; RNaseA; 1.  
DR SMART: SM00092; RNase_Pc; 1.  
DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.  
SQ SEQUENCE 155 AA; 17385 MW; 965CB118A5F5D5809 CRC64;  
  
OY Query Match 100.0%; Score 31; DB 11; Length 155;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 YTCR 5  
115 YTCR 119  
  
RESULT 13  
O9JRH8 PRELIMINARY; PRT; 155 AA.  
ID O9JRH8  
AC O9JRH8  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Eosinophil-associated ribonuclease 12.  
GN EAR12.  
OS Mus pahari (Shrew mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10093;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20243759; PubMed=10758160;  
RA Zhang J., Dyer K.D., Rosenberg H.F.;  
RT "Evolution of the rodent eosinophil-associated ribonuclease gene  
family by rapid gene sorting and positive selection.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).  
DR EMBL: AF238405; AAF67705.1; -.  
DR HSSP: P00656; 2RNS.  
DR InterPro: IPR001427; RNaseA.  
DR Pfam: PF000794; RNaseA; 1.  
DR PRINTS: PR00794; RIBONUCLEASE.  
DR ProDom: PD000535; RNaseA; 1.  
DR SMART: SM00092; RNase_Pc; 1.  
DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.  
SQ SEQUENCE 155 AA; 17167 MW; 651B846C9F485A35 CRC64;  
  
OY Query Match 100.0%; Score 31; DB 11; Length 155;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 YTCR 5  
115 YTCR 119  
  
RESULT 14  
O9JRH3 PRELIMINARY; PRT; 155 AA.  
ID O9JRH3  
AC O9JRH3  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Eosinophil-associated ribonuclease 12.  
GN EAR12.  
OS Mus caroli (Wild mouse) (Ricefield mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX NCBI_TaxID=10089;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20243759; PubMed=10758160;  
RA Zhang J., Dyer K.D., Rosenberg H.F.;  
RT "Evolution of the rodent eosinophil-associated ribonuclease gene  
family by rapid gene sorting and positive selection.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).  
DR EMBL: AF238411; AAF67710.1; -.  
DR HSSP: P00656; 1LSO.  
DR InterPro: IPR001427; RNaseA.  
DR Pfam: PF000794; RNaseA; 1.  
DR PRINTS: PR00794; RIBONUCLEASE.  
DR ProDom: PD000535; RNaseA; 1.  
DR SMART: SM00092; RNase_Pc; 1.  
DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.  
SQ SEQUENCE 155 AA; 17374 MW; BCB17A83A555495 CRC64;  
  
OY Query Match 100.0%; Score 31; DB 11; Length 155;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 YTCR 5  
115 YTCR 119  
  
RESULT 15  
O9JRG3 PRELIMINARY; PRT; 155 AA.  
ID O9JRG3  
AC O9JRG3  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Eosinophil-associated ribonuclease 32.  
GN EAR32.  
OS Mus caroli (Wild mouse) (Ricefield mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10089;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20243759; PubMed=10758160;  
RA Zhang J., Dyer K.D., Rosenberg H.F.;  
RT "Evolution of the rodent eosinophil-associated ribonuclease gene  
family by rapid gene sorting and positive selection.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).  
DR EMBL: AF238421; AAF67720.1; -.  
DR HSSP: P00656; 1LSO.  
DR InterPro: IPR001427; RNaseA.  
DR Pfam: PF000794; RNaseA; 1.  
DR PRINTS: PR00794; RIBONUCLEASE.  
DR ProDom: PD000535; RNaseA; 1.  
DR SMART: SM00092; RNase_Pc; 1.  
DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.  
SQ SEQUENCE 155 AA; 17188 MW; A83B912B6CCF8B39 CRC64;  
  
OY Query Match 100.0%; Score 31; DB 11; Length 155;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 YTCR 5  
115 YTCR 119  
  
Search completed: July 9, 2003, 12:02:03  
Job time : 17.75 secs
```

This Page Blank (uspto)